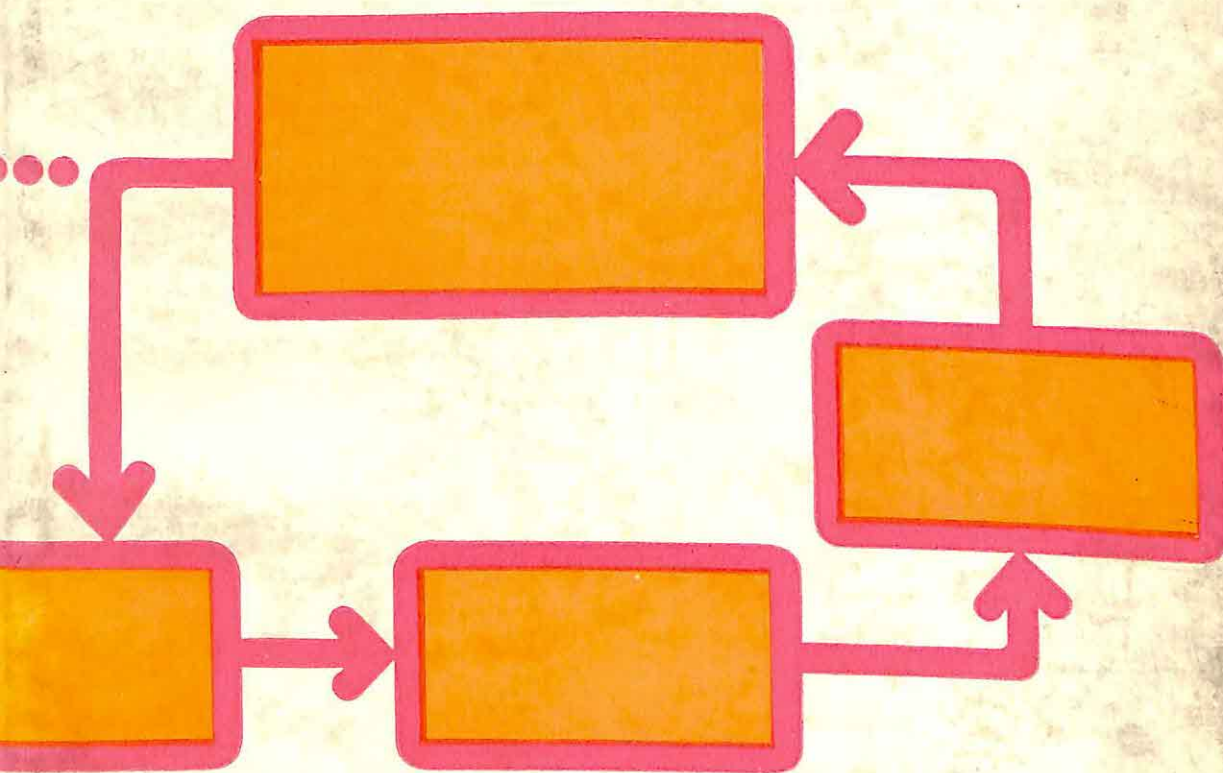


**Eastern
Economy
Edition**

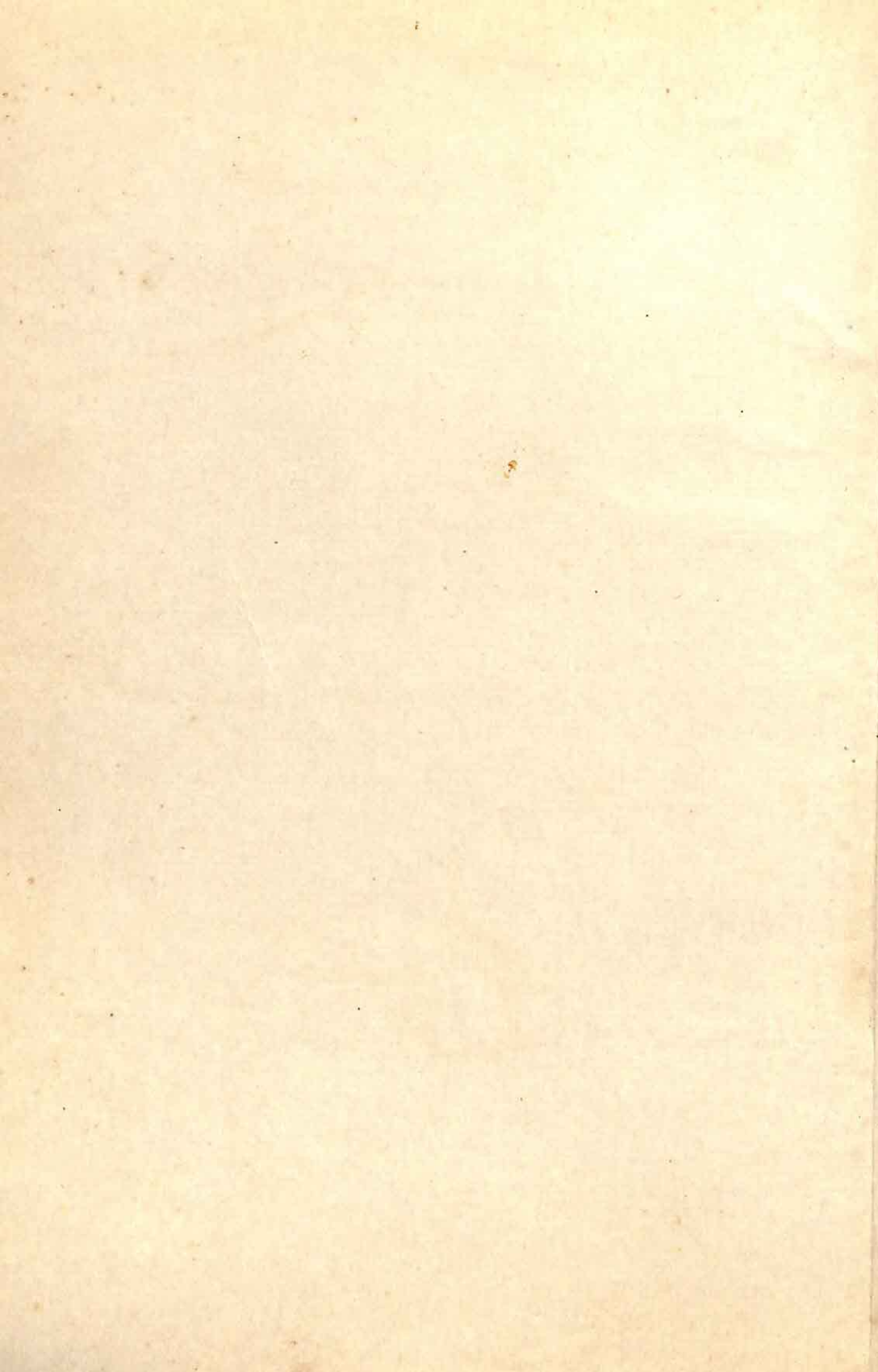
INVENTORY CONTROL:

theory and practice

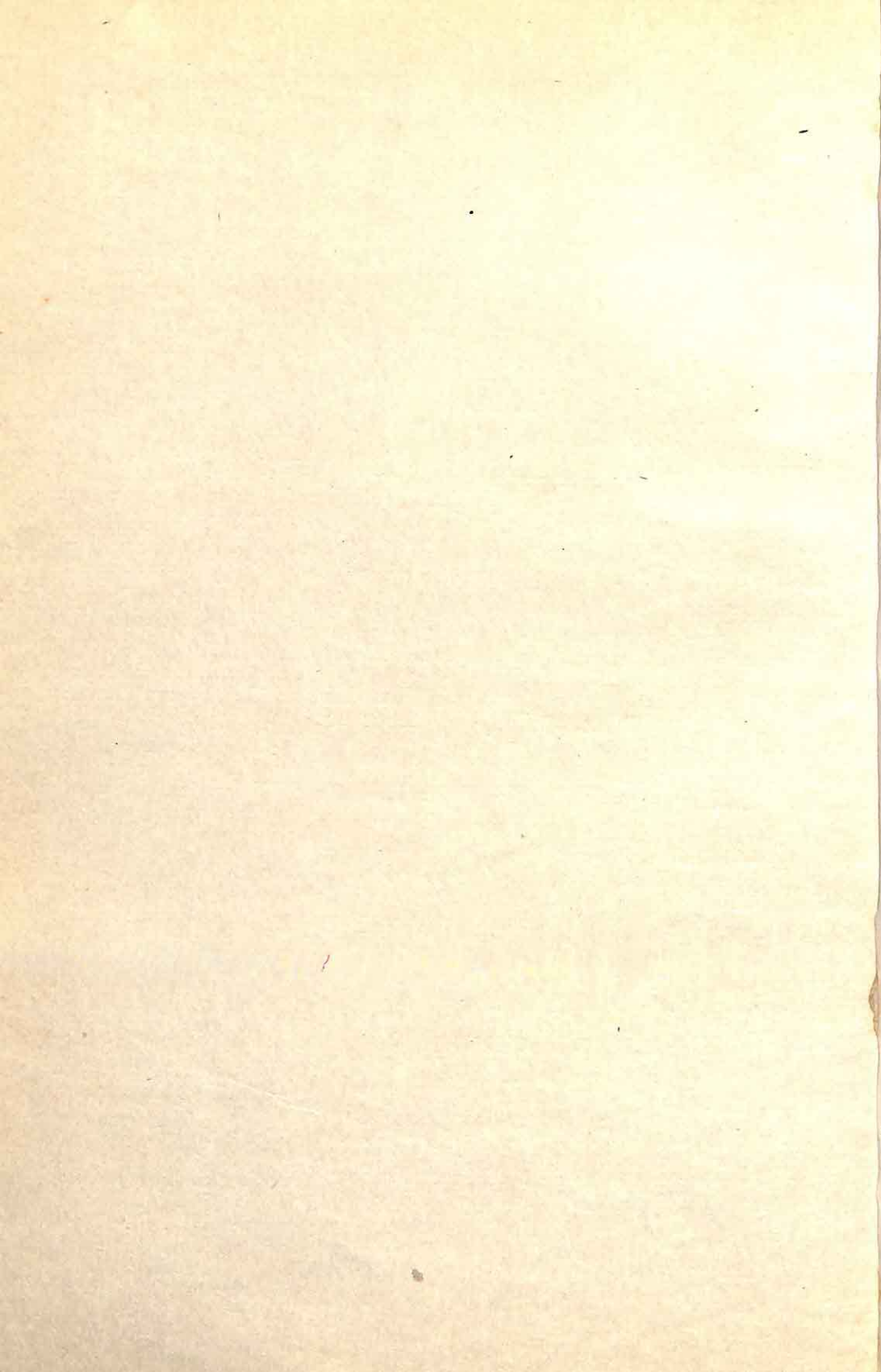


**martin k. starr and
david w. miller**

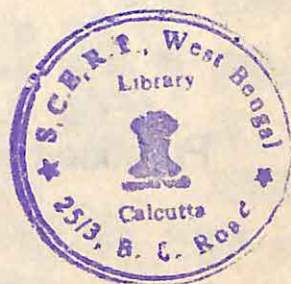




4938 15/3/91



inventory control:
theory and practice



C.E.R.T., West Bengal

ate 15.3.91

cc. No.

4938

658.787

STA

MARTIN K. STARR

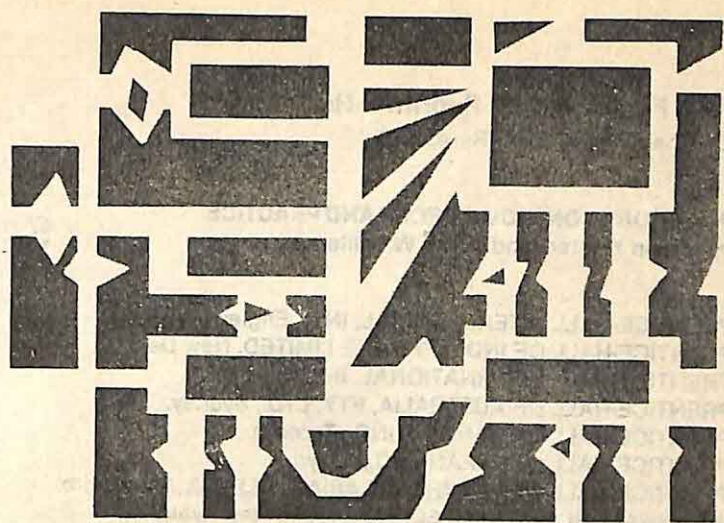
Associate Professors, Graduate School of Business, Columbia University

DAVID W. MILLER

Prentice-Hall of India Private Limited

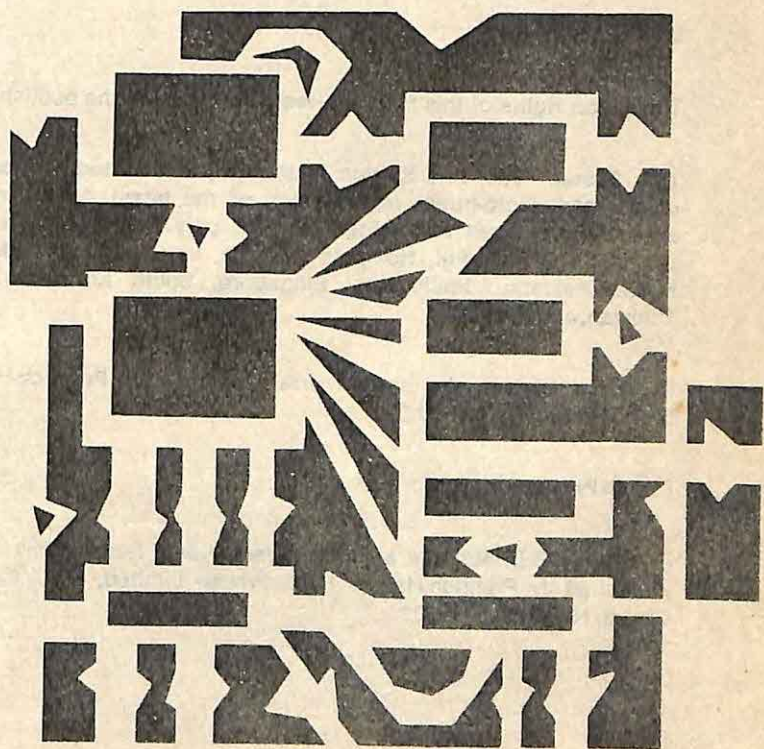
New Delhi-110001

1986



inventory control:

theory and practice



This Fourth Indian Reprint—Rs. 45.00

(Original U.S. Edition—Rs. 428.50)

INVENTORY CONTROL: THEORY AND PRACTICE

by Martin K. Starr and David W. Miller

658.787
STA

PRENTICE-HALL INTERNATIONAL, INC., Englewood Cliffs.
PRENTICE-HALL OF INDIA PRIVATE LIMITED, New Delhi.
PRENTICE-HALL INTERNATIONAL, INC., London.
PRENTICE-HALL OF AUSTRALIA, PTY. LTD., Sydney.
PRENTICE-HALL OF CANADA, INC., Toronto.
PRENTICE-HALL OF JAPAN, INC., Tokyo.
PRENTICE-HALL OF SOUTHEAST ASIA (PTE.) LTD., Singapore.
WHITEHALL BOOKS LIMITED, Wellington, New Zealand.
EDITORIA PRENTICE-HALL DO BRASIL LTDA., Rio de Janeiro.
PRENTICE-HALL HISPANOAMERICANA, S.A., Mexico City.

© 1962 by Prentice-Hall, Inc., Englewood Cliffs, N.J., U.S.A. All rights reserved. No part of this book may be reproduced in any form, by mimeograph or any other means, without permission in writing from the publishers.

ISBN-0-87692-184-5

The export rights of this book are vested solely with the publisher.

This Eastern Economy Edition is the only authorised, complete and unabridged photo-offset reproduction of the latest American edition specially published and priced for sale only in Bangladesh, Burma, Cambodia, China, Fiji, Hong Kong, India, Indonesia, Laos, Malaysia, Nepal, Pakistan, Philippines, Singapore, South Korea, Sri Lanka, Taiwan, and Vietnam.

Reprinted in India by special arrangement with Prentice-Hall, Inc., Englewood Cliffs, N.J., U.S.A.

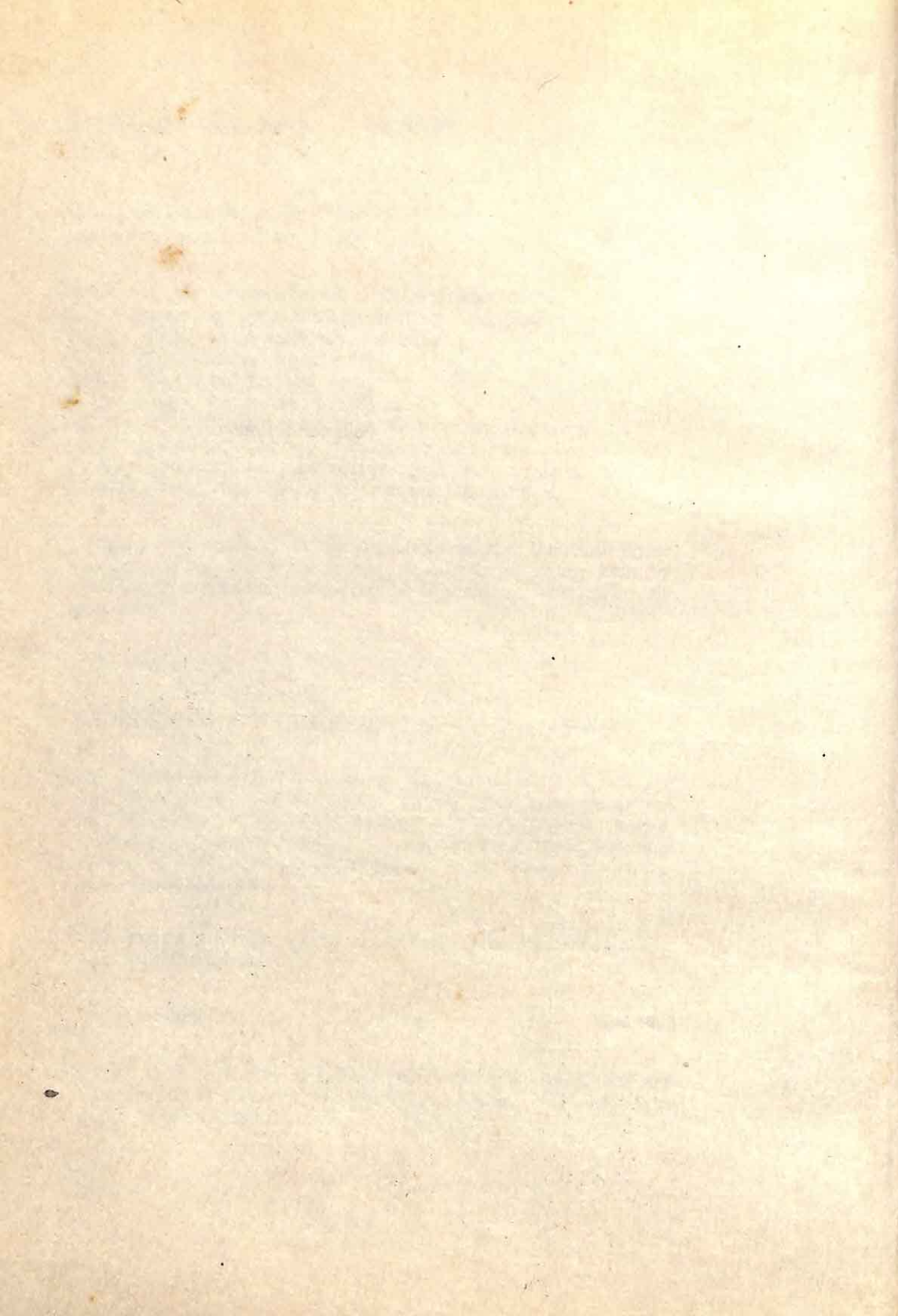
Fourth Printing

... ..
June, 1986

Printed by G.D. Makhija at India Offset Press, New Delhi-110064 and
Published by Prentice-Hall of India Private Limited, M-97, Connaught
Circus, New Delhi-110001.

to

Polly and Renée



preface

Most publishers supply author's manuals to their writers. One of the most important recommendations found in such manuals is that an author should constantly keep in mind the specific audience to whom his book is addressed. We have tried to observe this precautionary rule—the reader must judge whether or not we have succeeded. Since the preface is a suitable place to discuss such matters, we will indicate here some of the characteristics of the audience we have envisioned while writing this book. Our conception of the audience may help prospective readers to decide whether or not the book is directed to them.

Our subject is the control of inventories. In this field there is no need to apologize for adding another book to bookshelves already creaking in despair at the weight of volumes on the same subject. There are only a small handful of books devoted to our subject. This is in striking disproportion to the large periodical literature devoted to inventory theory and to the enormous number of successful applications of the theory to practical inventory problems. Further, there is no book on an intermediate level, and this is the level which we have aspired to achieve here.

What do we mean by an intermediate level? We can begin to answer this question by stating what we do *not* mean by "intermediate." This book is not a simple account of the marvels and magic of inventory theory, nor does it include accounts of the bonanzas which will result once the theory is applied. We assume that anyone having a passing acquaintance with the contemporary business scene already knows that there is good reason to utilize inventory theory. If he doesn't, this book is not intended for him. Good popularizations of inventory theory are already available and we have neither the desire nor the ability to produce a book competitive to them. At the other extreme, this book is not intended to be an illustration of the power of mathematics when it is applied to inventory theory. It so happens that inventory theory is mathematically interesting. By this we mean that mathematicians can find in our subject problems which exercise their ingenuity. However, there is no necessary connection between problems which are of interest to mathematicians and problems whose solutions have meaningful consequences in the practical applications of the theory. Our intention has been to confine consideration to the latter kind of problem, not the former.

The question of mathematics needs some separate comment. The authors of this book are not professional mathematicians. Accordingly, we have had to labor diligently to understand some of the more elegant mathematical papers which have appeared in the literature. Further, we too have noticed that there is no necessary correspondence between the promises made in prefaces with regard to mathematical level and the actual complexity of the book itself. It seems clear that it is not the level of mathematics which is used but the sophistication with which that level is used that causes communication difficulties. Since we are mathematically unsophisticated we hope that we have automatically avoided this particular problem. As to the mathematical level, we believe that a reasonable familiarity with calculus should enable one to comprehend this book. We will qualify our statement a little bit. An individual may be quite conversant with calculus and yet may never have had the occasion to take the derivative of a definite integral. Nonetheless, if the rule for so doing is known it is no more difficult than taking any other derivative. Since we use such derivatives frequently in inventory theory we have simply given the rule at an appropriate place in the text. Similarly, the use of Lagrangian multipliers may not have been covered in someone's calculus course, yet the procedure for using them is straightforward. We have therefore stated the basic procedure in the text and then have used the method as needed. These exemplify the exceptions to our statement that calculus is sufficient, and we hope they will only serve to prove the rule.

The preceding paragraph has suggested that we have had frequent occasion to use mathematical methods in the resolution of the inventory problems discussed in the text. This is, of course, required by the fact that, for most inventory problems, mathematical analysis is the only way to find the optimal inventory policy. Our intention has been to illustrate the major kinds of inventory problems and the kinds of analysis required to resolve them. We must emphasize the word "illustrate." There are such an enormous number of special cases in the general area of inventory problems that it would require an encyclopedia to discuss all of them. Therefore we have been forced to accept the more modest goal of "illustrating" the major kinds and types of inventory problems. This means that a specific example must be analyzed with enough completeness and generality to enable similar kinds of problems to be handled, even if they differ in details from the chosen example. Therefore, we have not hesitated to incorporate relatively lengthy discussions in the text. On the other hand, when some related topic seemed of lesser importance we have relegated it to our appendix. If the choice of what goes in the appendix sometimes seems arbitrary it is because it *is* arbitrary. It is mainly a question of taste—which, after all, is indisputable.

To return to the question of audience, this book is addressed to those who want an understanding of how different types of inventory problems can be solved. This interest may result from a need to actually solve an inventory problem. It might equally well result from the need to make decisions concerning the utilization of inventory theory in specific kinds of situations. We want to emphasize this second kind of motivation. The increasing utilization of computers for inventory control and the great importance of the systemic implications of any inventory control system demand that some attention be paid to the control aspects of the applications of inventory theory. The second part of the book is devoted to this subject. While the first six chapters of the book examine in detail the theory required for the solution of inventory problems, we have not followed the familiar pattern of devoting the rest of the book to practice. Instead, the subjects of the remaining chapters are closer to the theory of practice. They include methods for the analysis of the control aspects of different kinds of inventory systems and a variety of methods for estimating the advantages and disadvantages of various systems in given specific circumstances.

Finally, a word about the development of inventory theory. This wide-ranging theory has resulted from the labors of scores of persons: operations researchers, economists, mathematicians, engineers, and many others. Sometimes the specific contributions of particular individuals are treated in our presentation. When this is the case we have made appropriate references. More often the analyses are the result of many contributions and have, with time, gained a general anonymity. Obviously, our efforts to provide a comprehensive structure would have been impossible without the earlier contributions of these many others to whom the rich development of inventory theory should be credited.

M. K. S

D. W. M.

contents

part i: inventory theory

i. the analytical structure of inventory problems

1. on definitions, 3
2. the structure of inventory problems, 4
3. the structure of the analysis, 7
4. the relevant costs, 9
5. the measurement of costs, 13
6. the objectives of carrying inventory, 15
7. the plan of the book, 18

ii. static inventory problems under risk

8. general characteristics, 19
9. the christmas tree problem, 21
10. the opportunity cost matrix, 26
11. the cost of risk, 31
12. other cost structures, 33
13. the mathematical formulation: the discrete case, 36
14. the mathematical formulation: the continuous case, 40
15. imputation of costs, 44
16. size of work force, 45
17. general comments, 46

iii. static inventory problems under uncertainty

18. general characteristics, 49
19. decision criteria under uncertainty, 50
20. decision criteria for uncertainty and inventory problems, 53
21. objective and subjective probabilities, 55
22. partial information, 58
23. distribution-free inventory analysis: the Tchebycheff inequality, 62
24. distribution-free analysis: some other inequalities, 66
25. distribution-free analysis: per unit cost of understock, 72
26. comparison of analyses with full and partial information, 74

iv. dynamic models under certainty

- 27. general characteristics, 78
- 28. optimal lot-size model with constant demand, 79
- 29. quantity discounts, 84
- 30. optimal lot-size model with varying demand, 86
- 31. more than one item: the optimal policy curve, 93
- 32. multiple items from one supplier, 104
- 33. multiple warehouses, 110

v. dynamic inventory problems under risk

- 34. general characteristics, 112
- 35. the basic kinds of inventory control systems, 113
- 36. demand probability distributions, 116
- 37. the Q -system of inventory control, 120
- 38. the P -system of inventory control, 125
- 39. an alternative formulation of inventory policies: switching matrices, 130
- 40. markovian decision-making, 136
- 41. optimal selling policy with fluctuating prices, 142
- 42. variable lags between order and delivery: the queuing model, 146

vi. dynamic inventory problems under uncertainty

- 43. general characteristics, 152
- 44. the moments of convolutions, 153
- 45. Q -systems for inventory control, 155
- 46. P -systems for inventory control, 160
- 47. the use of central warehouses, 162

part ii: implementation phase: the theory of practice

vii. the evaluation study: design of the inventory study

- 48. elements of the inventory study, 172
- 49. intuitive and judgemental approach, 175
- 50. size of inventory investment and number of items carried, 176
- 51. variance of usage and item grouping, 181

52. controlling the reserve stock costs of the study, 190
53. controlling the forecast error costs of the study, 193
54. selecting the sample for the evaluation study, 196
55. exhaustion of study funds, 201
56. conclusion, 205

viii. the inventory study: design of decision procedures

57. systems analysis, 209
58. the many item effect, 211
59. the many location effect, 213
60. the many department effect, 214
61. a general discussion of flow diagrams, 216
62. a general discussion of input-output models, 218
63. testing alternative system designs by simulation, 221
64. monte carlo simulation—the Q -model, 227
65. detailed flow diagrams, 233
66. simulation of the P -model, 236
67. simulation of P -model with variable lead time, 239
68. simulation of inventory queuing model, 242
69. the many item simulation, 248
70. conclusion, 257

ix. the data processing study: design of operating procedures

71. the cost of an order, 260
72. perpetual vs. periodic data-processing, 262
73. capacity and load, 266
74. underloaded data-processing systems, 269
75. overloaded data-processing systems, 272
76. fluctuations in load, 274
77. regulation of data flow, 283
78. regulated inventory systems, 296
79. self-regulated inventory systems, 309
80. forecasting and updating controls, 316
81. conclusion, 326

appendix, 331

tables 337 ff

glossary, 340

bibliography, 344

index, 347

- 25 controlling the reactive stress effect on the brain
- 26 controlling the reactive stress effect on the brain
- 27 selecting the sample of the experimental study
- 28 definition of study limits
- 29 conclusion

VIII. the inventory study: design of decision procedures

- 30 inventory study
- 31 the inventory study
- 32 the inventory study
- 33 the inventory study
- 34 the inventory study
- 35 the inventory study
- 36 the inventory study
- 37 the inventory study
- 38 the inventory study
- 39 the inventory study
- 40 the inventory study
- 41 the inventory study
- 42 the inventory study
- 43 the inventory study
- 44 the inventory study
- 45 the inventory study
- 46 the inventory study
- 47 the inventory study
- 48 the inventory study
- 49 the inventory study
- 50 the inventory study

IX. the data processing: design of decision procedures

- 51 the data processing
- 52 the data processing
- 53 the data processing
- 54 the data processing
- 55 the data processing
- 56 the data processing
- 57 the data processing
- 58 the data processing
- 59 the data processing
- 60 the data processing
- 61 the data processing
- 62 the data processing
- 63 the data processing
- 64 the data processing
- 65 the data processing
- 66 the data processing
- 67 the data processing
- 68 the data processing
- 69 the data processing
- 70 the data processing
- 71 the data processing
- 72 the data processing
- 73 the data processing
- 74 the data processing
- 75 the data processing
- 76 the data processing
- 77 the data processing
- 78 the data processing
- 79 the data processing
- 80 the data processing
- 81 the data processing
- 82 the data processing
- 83 the data processing
- 84 the data processing
- 85 the data processing
- 86 the data processing
- 87 the data processing
- 88 the data processing
- 89 the data processing
- 90 the data processing
- 91 the data processing
- 92 the data processing
- 93 the data processing
- 94 the data processing
- 95 the data processing
- 96 the data processing
- 97 the data processing
- 98 the data processing
- 99 the data processing
- 100 the data processing

part one

inventory theory

the analytical structure of inventory problems

I. ON DEFINITIONS

Our subject is the body of mathematical and quantitative methods known as inventory theory. On first consideration the words, "inventory theory," seem to be virtually self-explanatory and a definition would appear to be obvious. "Theory" hardly requires elucidation and "inventory" immediately brings to mind a stock of some kind of physical commodity. The problem of the retailer, or the manufacturer, can be taken as a paradigm. In order to sell an item, he must maintain a stock of that item to meet the demand. As his stock is depleted he will order, or produce, some quantity of the item so that he can continue to meet the demand for it. Such being the nature of an inventory, it follows that inventory theory must deal with the logic which does, or should, underlie this procedure. A definition, then, might be: inventory theory deals with the determination of optimal procedures for procuring stocks of commodities to meet future demand. Recollection of the enormous variety of commodities for which someone must maintain stocks suggests that there would be a considerable number of special cases within inventory theory, so defined, but the definition appears to be broad enough to encompass all of them.

A little reflection suffices to indicate that perhaps we have not been sufficiently imaginative in our definition. As an example, let us consider currency—which is not a commodity. Supermarkets are requested to cash pay checks for their customers. At the start of any business day these requests represent a future demand for currency and the store manager must procure a stock of currency to meet the day's withdrawals. Extending the same idea a little further, it appears that companies have a similar problem

with regard to working capital. These problems seem to have the same fundamental characteristics as the more usual inventory problem on which we based our definition. It might be expected that inventory theory would have something to say concerning these problems as well.

It is certainly easy to modify our definition to include problems involving stocks of currency. All we need do is to relax our specification that inventory theory deals with stocks of commodities. Currency, while not a commodity, is certainly tangible and we can change our definition correspondingly. However, working capital is not quite as tangible as currency and we might well wish to include this kind of problem under our theory. It is easy to find analogous problems which deal with quite intangible things. In building a new plant it is necessary to choose a level of plant capacity to meet the future demand for the product. In signing a long-term lease one is selecting an amount of space to meet future demand for the space—whether it be a home or a business office. Department stores must provide an inventory of salesgirls' time to meet the future demand of customers. A railroad must determine the number of seats to provide for, say, the New York to Chicago run. The seats represent the supply available to meet the future demand—a typical inventory problem. Such examples could be multiplied indefinitely. All of them have in common that which we shall take to be the defining characteristic of the subject of inventory theory: the stocking of anything, whether tangible or not, to meet future demand.

This definition is sufficiently general, and sufficiently descriptive, to meet the needs of this book; however, one additional point should be made at this time. So far, nothing has been said about costs or values or, indeed, any economic characteristics at all. We will, of course, have occasion to discuss these aspects below. Nonetheless, it is clear that virtually all inventories, whether of tangible or intangible things, do have economic value. The recognition of this fact permits an alternative, but equivalent, definition which puts the problem of inventories in a different perspective. This definition was suggested by Fred Hanssman*: "An inventory is an idle resource of any kind, provided that such resource has economic value." Inventory theory, then, deals with the determination of the optimal level of such an idle resource. The element of futurity is still implied since an idle resource resulting from a past decision represents a sunk cost for future decisions. The planning in advance of the level of the idle resource is the subject of inventory theory in this definition.

2. THE STRUCTURE OF INVENTORY PROBLEMS

The bare bones of the typical inventory problem are already disclosed by our definition, but we can gain a better understanding if we add some flesh in the form of a further elaboration. This will serve to define better the inventory

* Fred Hanssman, "A Survey of Inventory Theory From the Operations Research Viewpoint," in *Progress in Operations Research*, ed. by Russell L. Ackoff (New York: John Wiley & Sons, Inc., 1961), p. 65.

problem itself and also will point out the natural divisions within the overall field of inventory theory. The multitude of special cases within inventory theory are a direct result of the differences among practical inventory problems. These differences, in turn, stem from variations in the major structural components of the general inventory problem. This section will be devoted to a consideration of this structure.

Our definition calls attention to two aspects of the inventory problem: procurement of the commodity or thing in question and the future demand. Each of these introduces some important differentiations among inventory problems. Let us consider demand first. Obviously, some specific level of demand will eventuate at any given point in time. But at the time the inventory decision must be reached the important question is what do we know about the level of future demand. Following traditional practice in decision theory, we can conveniently summarize the possibilities with regard to our knowledge of future demand in three categories. First, we may know exactly what the future demand will be. This is not a very frequent case but it may be illustrated by the construction of a skyscraper with regard to the inventory of steel girders. In such a case we might well know exactly how many girders we would need during each week of the construction. Such a case is called an inventory problem under certainty. Second, we may know the probability distribution of future demand. Such information is likely to be available if the item in question is one for which records of past demand are available. An example might be the inventory of tires for a taxicab fleet or the inventory of bread in a supermarket. This case will be called an inventory problem under risk. Third, we may be entirely ignorant of the likelihood of various levels of future demand. Utter ignorance is probably just as rare as complete certainty and we shall consider ways in which partial information can be used, but for present purposes it is useful to introduce this extreme case as one possibility. An example might be the inventory problem of plant capacity for a new product for which there are no existing market analogies. This case will be called an inventory problem under uncertainty. These, then, are the three essentially different kinds of cases of inventory problems with regard to our knowledge of future demand.

The procurement process introduces two kinds of differentiations. Generally, there is some time lag between the time when an order is placed and the time when the commodity or thing in question is actually received in inventory. The first difference arises in this regard. For some kinds of inventory problems this time lag is constant or essentially constant. For other inventory problems there is a probability distribution of possible time lags. This distinction has important consequences in the analysis of the inventory problem. The second difference results from the fact that some companies order the commodity in question from an outside supplier while other companies produce the commodity themselves. In the latter case, in effect, one part of the company orders the commodity from another part of the company. This difference has a major effect on the analysis of the inventory problem because the self-supplying company must consider the effect of its ordering policies not only on the inventory situation but also on

the production process. In other words, the inventory problem of the self-supplying company is likely to be considerably more complex than that of the company which is supplied from the outside.

A third important difference between various inventory problems is not directly suggested by our definition. This difference arises because some inventory decisions are essentially made only once whereas other inventory decisions are simply steps in a continuing process of such decisions. Examples of both kinds are easy to find. A decision concerning plant capacity must be made when the plant is first planned. While plant expansions can be subsequently undertaken, or further plants can be built, still any such decision is likely to be sufficiently distant in time to justify considering each such decision as an independent one. For some kinds of fashion goods it is impossible to obtain replacement stock during the given season. Therefore, the total inventory desired must be obtained on the first order. These are examples of a one-shot inventory decision. On the other hand, most supermarket and department store inventory problems involve repeated orders for the same commodity and, hence, are of the nature of a continuing process of inventory decisions. This difference should not be understood in any absolutistic sense. Virtually any inventory decision could be conceptualized as one in a series of similar decisions if a sufficiently long time span were considered. Nonetheless, the distinction remains an important one. In the continuing process the outcome of one decision has an immediate effect on the subsequent decisions and it is necessary to take this effect into account in the analysis. This requires, then, a kind of dynamic analysis. In the one-shot decision it is not necessary to consider the effect of the decision on subsequent decisions. This permits the utilization of static kinds of analysis, with concomitant simplification as compared to the dynamic case.

Still another distinction is important in the case of the dynamic, repetitive kind of inventory decision process. This results from the fact that for some inventory problems of this type the demand level is fixed over time, typically in the form of a constant probability distribution of demand. In other inventory problems this is not the case and the probability distribution of demand varies over time. This differentiation is important primarily because of our analytical shortcomings. The mathematical complexities of the analysis in the case where the demand distribution varies with time are such that it is much more difficult to obtain readily usable results. We are generally on much safer ground when we can assume a fixed demand distribution over time.

These, then, are the most important differences among inventory problems. Let us summarize them:

- (1) Knowledge of demand: certainty, risk, uncertainty.
- (2) Method of obtaining commodity: outside supplier or self-supplying.
- (3) The decision process: one-shot decision = static, or repetitive = dynamic.

These are the most important differentiations. We have also noted:

- (4) Analytical convenience in the dynamic case: fixed demand distribution over time or varying demand distribution.

(5) Time lag in receiving order: constant time lag or probability distribution.

There are numerous other distinctions which lead to a host of special cases. Some of them might have been included in our list. However, we think the other distinctions are of lesser importance, with a major exception to be noted in the next paragraph. Such differences result from, for example, multiple warehouses with and without central ordering, cascades of inventories from raw materials through in-process inventories to finished goods inventories, multiple-item orders from one supplier, and so forth. In our subsequent discussion and analysis we will have occasion to consider some of these special cases but at this point it suffices merely to note their existence.

One most important difference among various inventory problems has not even been mentioned yet. This is the question of the relevant costs which will enter into our analysis of any specific inventory problem. However, the discussion of costs is of such crucial importance to the whole subject that it deserves separate treatment and we will return to it shortly.

3. THE STRUCTURE OF THE ANALYSIS

The analysis of inventory problems is fundamentally based on a very simple, common-sensical observation. This is that in any genuine inventory problem whatsoever there *must* be opposing costs. By this we mean simply that there must be a cost associated with doing "too much" and there must be a cost associated with doing "too little." Sometimes there are several such costs but there must always be at least one in each direction. These statements are quite vague and the immediate question is: "too much" or "too little" of what?

This question can be answered most directly by a slight anticipation of some subsequent discussion. The resolution of any specific inventory problem requires the answers to two questions. First, when or how often should the commodity be ordered? Second, how much of the commodity should be ordered on any particular order? Not every inventory problem demands an answer to both questions. For example, in a one-shot inventory decision we know, by definition, that there will only be one order. However, the general case will require the answers to both questions. Furthermore, when these questions are answered, the inventory problem is solved. The inventory decision problem consists, precisely, of determining these two answers. This granted, let us consider the two questions separately in terms of what we have called opposing costs.

How often should the commodity be ordered? The precise answer requires analysis of the specific problem but we know two things about the frequency of order. First, there must be a cost associated with ordering too frequently. If this were not the case then the commodity would be ordered with the maximum possible frequency, perhaps a separate order for each item demanded. Now, as a matter of fact, this may well be the optimal ordering policy for some commodities but this is not the point at issue. If there were

no cost associated with ordering too frequently then there would simply be no problem and no analysis would be required. It is precisely the existence of such a cost which creates the problem. Second, there must be a cost associated with not ordering frequently enough. If there were not such a cost the commodity would never be ordered at all or, alternatively, one order might be placed for a gigantic amount of the commodity in question. If either one of these opposing costs were non-existent there would be no inventory problem with regard to ordering frequency and the optimal policy would be to go to one of the two indicated extreme solutions.

How much of the commodity should be ordered? Once again we can be sure that our two opposing costs must exist. If there were no cost associated with ordering too much then some enormous quantity would automatically be ordered. Correspondingly, if there were no cost associated with ordering too little then no stock would be kept in inventory. If these costs did not exist there would simply be no inventory problem in terms of the amount of the commodity to be ordered. So here, also, we find the opposing costs we referred to in the first paragraph of this section. In every inventory problem there are such opposing costs and the first step of the analysis must be to determine what the costs are and then, if possible, to measure them. We will discuss these questions at some length in the next section.

Granted the existence of such opposing costs, how does the analysis proceed? Basically, by recognition of the fact that the usual objective in inventory problems will be the minimization of the total costs involved, not simply a minimization of one or the other of the opposing costs. The total costs include all of the relevant opposing costs and the analytical procedure is to find an expression for the total costs and then to discover the course of action which will minimize them. The method used in accomplishing this depends directly on the state of our knowledge concerning future demand. If the demand is known with certainty, and assuming that the relevant costs can be measured, the total costs can be determined for each possible course of action and that one which has the lowest total cost can be directly selected. If we are dealing with an inventory problem under risk we cannot know in advance which of the various possible levels of demand will in fact occur. However, since we know the probabilities with which the various levels of demand will occur we can calculate the expected cost for each possible course of action. Then we can select that course of action which minimizes the expected total cost. If the inventory problem is under uncertainty there is no generally accepted procedure to be followed but there are some eminently reasonable approaches which can be used. We will reserve discussion of this case for subsequent, separate treatment.

The preceding paragraph tacitly assumes that the intention of the analysis of an inventory problem is to determine the absolutely best course of action. This, of course, is the best of all possible analytical goals and whenever it can be accomplished it will be. However, we should note explicitly that this is not always the actual goal of our analysis. Sometimes the inventory problem is overly complex or, as we shall see, we may not be able to obtain all the information necessary. In either of these cases we may well be satisfied with

the lesser goal of improving the existing situation without concerning ourselves with the more difficult problem of determining the optimal course of action. This situation often arises in practice and it frequently happens that very large savings can be accomplished without any necessity of having to determine the true optimal course of action.

4. THE RELEVANT COSTS

Our preceding discussion has centered attention on the crucial importance of the determination and measurement of the relevant costs in the analytical resolution of inventory problems. We must now consider these costs in more detail. What is the nature of them? How can we measure them? We will try to discuss the major classes of inventory costs. Each specific cost arises in some kinds of inventory problems but generally not all of the various costs will be simultaneously relevant to any single inventory problem.

The first major class of costs is the procurement costs. It is customary to distinguish between procurement costs when outside suppliers are involved, the ordering cost, and the same kind of cost when the commodity is self-supplied, the set-up cost. These costs, in both cases, play the same role in the analytical formulation of the inventory problem. The ordering cost includes all those cost components which result from the processing of an order. In order to send out an order it is necessary to review the given item and determine how much must be ordered. Then the order must be clerically processed, which typically requires contributions of time from several individuals. Finally, it will usually be true that each order will require a separate check for payment and this requires further clerical processing. It is often possible to determine the variable costs per order directly from cost accounting data. However, it must be remembered that costs in the accounting sense are not always the costs we need for the analysis of a decision problem. This same question arises for all the costs which we will need to consider so we may as well take a paragraph to indicate the nature of the difficulty involved.

The difficulty can be formulated in two essentially equivalent ways. The first way is based on the common distinction between fixed and variable costs. The economic questions revolving around this distinction are amply illustrated by John Maurice Clark's classic work, *Studies in the Economics of Overhead Costs*.^{*} Similar problems arise in the analysis of decisions. We can offer a completely satisfactory definition of fixed costs in terms of any decision problem: those, and only those, costs which do not vary for any available course of action in any specific decision problem are fixed costs for that decision problem. Such fixed costs are sunk costs for the decision problem in question and can be ignored. The difficulty arises, then, because accounting procedures are constructed with the intention, so to speak, of providing a kind of average fixed costs for the whole array of managerial decisions.

^{*} J. Maurice Clark, *Studies in the Economics of Overhead Costs* (Chicago: The University of Chicago Press, 1923).

For any specific decision the accounting fixed costs may not be at all correct. For example, suppose a company has an order department with five clerks and a separate office. Suppose, further, that the inventory policies can be changed so that ordering can be accomplished by only two clerks in a smaller office. From the standpoint of accounting data the wages of the clerks would be variable costs and the ordering department overhead burden, the fixed costs, would be determined in accordance with the space utilized. From the point of view of the inventory decision this may or may not be correct. If, for example, the clerks, because of union regulations or any other reason, could not be fired and could not be utilized elsewhere in the company then the cost of their wages is fixed, and sunk, for the inventory decision problem. Similarly, if some other utilization of the space were available then the "fixed" overhead cost is a variable cost for the decision problem. The second way of formulating the difficulty is in terms of the distinction between costs and opportunity costs. This distinction, however, is so important that we will reserve discussion of it until we are able to offer some numerical illustrations.

The self-supplier's procurement costs are called set-up costs. This name is, strictly speaking, only correct when we are considering the case of a company with a production line which makes a number of items on a job-order basis. Set-up cost then refers to the cost of changing over the production process to produce the ordered item. This always entails time lost for the production process and, thus, involves an associated cost which can usually be determined directly from cost accounting records. In inventory problems we add to this set-up cost of the production process any clerical costs involved in sending an order to the production department. Thus, set-up cost includes all of the cost components associated with one order for the given item. The case of the company with a continuous production line is somewhat different. Such a company's inventory problem is, by definition, of the kind we have called dynamic. The inventory decision problem is the same as the production scheduling decision problem for such a company. We will consider this case subsequently and here we will simply note that the analogous cost in this case is the cost which results from varying the production level—the costs of hiring and firing, training, and so forth.

The second class of costs we will call the stockage costs: the costs of carrying and of not carrying inventory. The first such cost is the carrying cost itself. This includes a number of component costs, not all of which will necessarily be involved in any specific inventory problem. Let us describe some of these component costs:

(1) *The cost of the money tied up in the inventory.* The amount of money invested in the inventory could be utilized elsewhere to earn some kind of return. Since it is tied up in inventory it is not available and this fact requires that a cost be assigned to reflect lost earning power. The cost that should be assigned depends on the use to which the money would be put if it were available. Thus, a company might be earning a return of 10% on its total investment. If money released from inventory could also earn 10% then this is the cost of the money tied up in inventory. Alternatively, many

companies maintain short-term loans from banks. Money released from inventory could be used to retire some of these loans. In this case, the cost of the money in inventory might be measured by the interest rate paid for the short-term loans. Also, the money might be used to retire company bonds in which case the interest rate paid on the bonds would be used. If none of these apply, at the very least the company could invest the money in government securities and earn the going rate on them. This rate, then, could be used as the cost of the money tied up in inventory.

(2) *Storage costs.* The space required to store the inventory of the given item usually has a cost associated with it. We say "usually" because this, too, depends on whether there is an alternative use for the space in question. If there is not—for example, if a warehouse can not be sold or rented and there is no alternative profitable use for the space within the company—then the space is a fixed cost for the decision concerning inventory.

(3) *Deterioration costs.* Many kinds of commodities and items deteriorate in value during storage. This can result from actual deterioration, obsolescence, or even pilferage. This loss in value, then, represents a cost which must be assigned to carrying inventory.

(4) *Insurance costs.* Since many inventories require insurance it is necessary to include this cost in the cost of carrying inventory. Note that this is equally true whether outside insurance is carried or the inventories are self-insured.

A second kind of cost associated with carrying inventory results when there is stock left on hand after the demand for the item has terminated. This cost is called the overstock cost. The interpretation of this cost differs in accordance with whether the inventory problem is static or dynamic. Consider a simple, and well-known, example of a static inventory problem: that of a merchant who wishes to sell Christmas trees. His sales season is very sharply limited and any trees left on hand at the end of the season will have very little, if any, salvage value. Thus, if he has too much stock he will have a cost due to the loss he will suffer for each tree he overstocks. This cost would be the overstock cost for this inventory problem. For dynamic inventory problems there are two kinds of cases. Consider, as one example, the inventory problem of a department store with regard to some standard household item such as towels. There is always a demand for towels so any towels left over after one time period can be sold in the subsequent periods and this remains true for an indefinite period in the future. Therefore, there will never be an overstock in the sense above and, hence, no overstock cost. Note that the carrying costs for this problem will already have been taken into account. For the second example consider the problem of inventories of high-fashion items. In women's clothes there are two selling seasons, Fall and Spring. If only one order can be placed we have the static inventory problem already mentioned. If several orders can be placed during a season then we have a dynamic problem. There will be no overstock cost involved until the last order period of the season. At that time, however, any overstock will typically undergo a drastic devaluation and this must be measured by the overstock cost.

The cost of not carrying inventory is called the out-of-stock cost. There are two variants of this cost, depending on the reaction of the prospective customer to the out-of-stock situation. One kind of case occurs, for example, in mail-order sales. If the company is out of stock when an order is received it will typically institute an emergency expediting procedure to get some stock. This situation is called a back-order. The sale to the customer will not be lost, only a delay of a few days in shipment. However, as a result of the out-of-stock situation there will be additional costs: the cost of expediting, special handling costs, and often special shipping and packaging costs. All of these, then, must be included in the out-of-stock cost. The other kind of case occurs when the sale is lost. For example, it is often the case that a prospective customer in a phonograph record store refuses to place an order for an out-of-stock item. Instead, he simply goes to a competitive store which has the item in stock. Ordinarily, this must be considered to result in a cost; the loss of goodwill is the usual rubric which describes it. It is tempting to think that this cost might be measured by the profit lost due to not making a sale but this is incorrect for two reasons. First, it is an out-and-out error to measure the cost in this way because it amounts to a confusion of ordinary costs with opportunity costs. We will illustrate this below. Second, when we refer to the cost due to the loss of goodwill we are thinking of the introspectively plausible, and behavioristically verifiable, fact that a prospective customer who has found one or more out-of-stocks in a given outlet will become less likely to return to that outlet for other purchases. Therefore, the store suffers a loss of future sales which might otherwise have been made. It is this cost which we wish to include in our formulation. Let us hasten to add that it is extremely difficult and often virtually impossible to measure this cost. When this is the case it is usually necessary to rely on executive estimates of the cost. Another kind of out-of-stock cost results when a self-supplier inventory problem is under consideration. Typically, the effect of an out-of-stock situation under these circumstances will range from the necessity of substituting an inferior, or more expensive,* item for the out-of-stock item to the interruption of some production process because of the lack of a crucial item. For example, if a company runs out of stock of some bolt it uses in its production process it may well be able to use a similar bolt of some other material or a more expensive kind of bolt. The additional cost undergone because of this substitution is the out-of-stock cost for the bolt in question. Alternatively, if a company runs out of some crucial raw material the whole production process may have to shut down, with all the large costs attendant on such a happening. This would be the out-of-stock cost for the given raw material in this case. It can be seen from these examples that the out-of-stock cost is sometimes a cost per unit out-of-stock, as in the case of the bolts, and sometimes is a fixed amount (which may be a function of time), as in the case of the raw material.

The third, and last, class of costs we shall call the systemic costs. These

* This policy may also be followed by the mail-order house in our preceding example.

costs have not been emphasized in the literature on inventory theory but they are of considerable importance in practical applications of the theory. Systemic costs generally enter into the formulation of the inventory problem when it is of the sort we have called dynamic. We can not adequately discuss these costs until we begin our analysis of dynamic inventory problems. Here we are satisfied to call attention to the existence of this class of costs and to give a very rough idea of their origin and nature. As we shall see subsequently, there are several alternative inventory control systems which might be used in the case of dynamic inventory problems. These systems differ in the amount of inspection of existing inventory levels which is required and the amount of data-processing necessary. They also differ with regard to the flexibility of the systems in terms of achieving potential savings which might result from the amalgamation of orders for several items into one order from the supplier. There are other potential savings which might result from amalgamation of orders from various warehouses, for example, into one central order. Further, there may be large differences in installation and implementation costs. Systemic costs are intended to cover these kinds of costs. We shall discuss them more extensively in our section on dynamic inventory problems, and in Part 2, which is primarily concerned with systemic factors.

These, then, are the major costs which enter into the formulation of an inventory decision problem. Not all of them are likely to be involved in any one inventory problem but the total cost expression for any specific inventory problem will be composed of the sum of some of these costs. The ability to measure the costs is of crucial importance and we must devote some further attention to the problems of measurement.

5. THE MEASUREMENT OF COSTS

In the preceding section there are a few scattered remarks concerning the possibilities of measuring specific costs. For several reasons we do not propose to enter upon an exhaustive discussion of the procedures which might be followed in attempting to measure the costs. First, such a discussion would have to be virtually book length by itself. It would, for example, have to include a lengthy section on cost accounting techniques. Second, there is no general approach to the problem such that when we apply it we can be assured that we will be achieving correct measurements of costs. On the contrary, in practice we are forced to utilize whatever ad hoc methods can be discovered which seem applicable. Third, the resolution of the cost measurement problem depends very much on the kinds of company records which are available and virtually no two companies are alike in this regard. Instead of attempting such a discussion we want only to make some general comments which are relevant to the problem.

In practice, some of the costs may be adequately determined directly from cost accounting records, at least in the case of the orthodox commodity applications of inventory theory. This requires due attention to the fact,

noted above, that the accountant's classification into fixed and variable costs will often not correspond to the one we need. In some cases—an example is the determination of the cost of varying production level—we can use the methods of regression analysis in conjunction with historical cost records. In other cases we may have to rely on executive opinion. We cited the cost of loss of goodwill as an example of this. How might it be done? One approach would be to simply ask the relevant executive what he would pay to have avoided one out-of-stock, called a back-order, for the item in question. His estimate may not be "correct" but, if nothing better is available, it at least has the merit that the resulting inventory decisions would be in accord with the opinions of the executive. Generally it is possible to measure all the costs involved in the usual kind of inventory situation with sufficient accuracy to achieve a resolution of the inventory decision problem. We say "sufficient accuracy" because, as we shall see below, the optimal course of action is not likely to be very much changed by even relatively large errors in the measurement of costs. In other words, the typical inventory analysis is not overly sensitive to reasonable errors in the measurement of costs. This fact helps to reassure us that our conclusions will be essentially correct even if we are unable to obtain precise cost measurements.

In our discussion of the definition of inventory theory we pointed out that the theory should be applicable to various problems which involved the "storage" of intangibles. We must now chasten the hopes that this might arouse by pointing out that the more "intangible" the subject of the inventory analysis, the more difficult it is to achieve measurements of the relevant costs. It is the problem of measuring costs which prevents the ready application of our theory to many of these very important kinds of problems. How, for example, can one measure the out-of-stock cost for a bank trying to use inventory theory to determine the optimal amount of cash to have on hand? Many companies have the "inventory" problem of maintaining a sufficient number of junior executives to meet the demand resulting from the retirement of senior executives. How can the costs be measured for this problem? A research program might be conceptualized as an inventory problem. The company wants to maintain an inventory of new products or product quality improvements against the demand resulting from competitive actions and other kinds of market changes. How can the relevant costs be measured? The problems of measuring such costs as these are truly enormous. We have noted the imperfections of cost accounting classifications from the point of view of the usual kind of inventory analysis. Nonetheless, it is precisely in those areas where cost accounting information is most readily available that inventory theory has been most successfully applied. Many of the problems which have the same formal structure as the everyday, garden variety of inventory problem but for which almost no accounting information is available will not be analytically resolved until we have discovered some new ways to measure costs.

* See Holt, Modigliani, Muth, and Simon, *Planning Production, Inventories, and Work Force* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1961).

If the preceding paragraph is correct, why did we insist on the relevance of inventory theory to some of these problems? For a very good, and most important, reason. The basic structure of the analysis would remain the same for these more intangible kinds of inventory problems if we knew the costs. The usual procedure is to reason from the costs to determine the optimal course of action. However, it also works in reverse. Any course of action which is selected implies some assumptions about the costs. Specifically, any selected course of action imputes a value to the ratio of the opposing costs in question. We shall illustrate this fact subsequently. For the moment we want only to emphasize the implications thereof.

In effect, this means that the estimation of the costs involved cannot be avoided. It is very easy to reason like this: "Well, if it is so difficult to get correct measurements of the costs I am simply going to ignore the problem—I will use my experience and good common sense to choose a reasonable inventory policy; I'll skip the theoretical analysis!" Unfortunately, this reasoning won't work. Whatever course of action our friend selects on the basis of his experience and common sense will inevitably impute a value to the ratio of the opposing costs. The only question is: does he want to make his estimate explicit or is he satisfied to leave it implicit and unknown to him? When the question is put this way it seems clear that, whatever the difficulties attendant on the cost measurement process, we are better advised to use whatever explicit information we can get than simply to ignore it.

This argument provides a very useful tool to assist the executive even in cases where the costs cannot be determined. Suppose, for example, that we are analyzing the inventory problem for a specific item. Assume that we have determined that the carrying charges are \$5 per month per item in stock. Assume, further, that the opposing cost in question is the loss of customer goodwill due to out-of-stocks and that we cannot obtain an estimate of this cost. We can take the existing policy and determine what cost it imputes to the loss of customer goodwill. We may find that the imputed cost under the existing policy is \$25,000—this order of magnitude is quite commonly found—and we can report this fact to the executive. One can imagine the dismayed expostulations which might result and which would precede his order to change the policy. The point is that we can use this idea of the imputation of the cost ratio to provide a valuable aid to the executive in using his experience. This same approach is available in the case of the intangible inventory problems and this is why we have emphasized the applicability of inventory theory to these kinds of problems even though it is quite unlikely that we can successfully measure the relevant costs.

6. THE OBJECTIVES OF CARRYING INVENTORY

All of our discussion until now has tacitly assumed that costs and/or profits were a completely satisfactory measure of payoff for the analysis of inventory problems. In other words, we have assumed that the company objectives could be satisfactorily measured, at least as far as inventory decisions are

concerned, in terms of dollars. Now, many executive decision problems other than ones dealing with inventory can be formulated in quantitative form and it is notoriously the case that dollars are often not suitable measures of the degree of achievement of the objectives involved.* Might this not also be true in the case of inventory decision problems?

It seems clear that the answer to this question is in the affirmative. For some inventory problems dollars of cost or profit might not be a satisfactory measure of payoff. However, it also seems quite true that the overwhelming majority of practical inventory problems are of a kind for which dollars are a perfectly realistic measure of payoff. There is far less likelihood that a different measure would be needed for inventory problems than for other classes of decision problems. We are going to use these statements to justify a concentration of our attention on inventory problems for which dollars represent an adequate measure of payoff. Therefore, it will be desirable to offer a little more discussion of them.

First, in any decision problem—inventory or not—when would dollars not be a satisfactory measure of payoff? There are two major ways in which this can happen. The first major reason is that the objective of the decision-maker may simply not be expressible in dollars. Such is the case, for example, when a company objective is to achieve a maximum possible share of the market or when a decision-maker's personal objective is to minimize the number of decisions he has to make. This kind of situation is unlikely to arise in inventory decision problems for the reason that inventories are not ends, or objectives, in themselves. The typical case is one in which we wish to meet some level of demand or of self-use and we want to do so as economically as possible. In other words, the utilization of dollars as a measure of payoff is built into most inventory problems. The second major reason for dollars being an unsuitable measure of payoff in the general decision problem is that the decision-maker's utility for dollars may not be measured by the dollar amounts. This will typically be the case when a considerable proportion of the decision-maker's resources are at risk in the given decision problem. Pierre Massé† has suggested that there are three classes of cases:

(1) The amount at risk represents only a small percentage of total resources. In this case dollars are a suitable measure of payoff.

(2) The amount at risk represents a sufficiently large percentage (perhaps 5%) to constitute a serious depletion of resources in the event of bad luck. In this case some measure of utility is needed, perhaps the logarithm of the amount in dollars.

(3) The amount at risk is so large that its loss would be equivalent to bankruptcy. In this case Massé suggests that the only suitable measure of utility is the probability of bankruptcy itself.

These three cases are not to be understood in terms of expected values.

* For examples see Miller and Starr, *Executive Decisions and Operations Research* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1960).

† Pierre Massé, *Les Reserves a la Regulation de l'Avenir dans la vie Economique*, 2 vols (Paris: Hermann, 1946).

The question is whether any specific course of action has any consequences which might entail losses of the indicated orders of magnitude. Since one can always find some remote possibility of large losses for any course of action it is important to add to Massé's classification a qualification suggested by Borel.* This is that in human affairs there are certain levels of probabilities which are small enough that they can be ignored—we simply assume that the corresponding events are impossible. Borel reasons from the probabilities of events which we ignore in our planning—such as being struck with lightning—to deduce that probabilities of less than one in a million (.000001) can be ignored. Now, in terms of inventory decision problems we can certainly think of cases of types 2 and 3 of Massé's classification. However, it is certainly true that the vast majority of practical inventory problems involve only marginal percentages of resources at risk. It is for this reason that we will not concern ourselves very much with inventory problems other than those for which we can satisfactorily measure the payoffs in dollars of cost or profit.

We noted that inventories are typically not ends in themselves. What, then, are the major purposes which they serve? Kenneth Arrow† classifies the purposes into three kinds, suggested by Keynes as the motives for holding stocks of cash: transaction, precautionary, and speculative motives. The transaction motive results from the fact that it is not generally possible even in the case of certainty to synchronize perfectly the inflow and outflow of the commodity in question. Inventories are therefore held in order to compensate for the lack of synchronization. The precautionary motive results from the usual inability to predict demand exactly—most inventory problems are under risk—and the consequent need to maintain some kind of a safety allowance. We may note that this motive will only operate as a result of the inability to obtain instantaneous delivery of commodities, at least without extra costs. The speculative motive results when prices are rising or when there are expected changes in costs. Under these circumstances profits may be made by holding inventories at the lower price until the higher price obtains.

In our analysis we will be utilizing essentially this three-fold breakdown. The transaction and precautionary motives are basic to our analytical framework and we will introduce the speculative motive when it seems appropriate. It should be noted that for the usual inventory problem the speculative motive is of lesser importance than the other two. Many companies, as a matter of fact, establish policies specifically to prevent the operation of this motive. For example, consider a company which utilizes soybeans as a raw material. The typically fluctuating prices of soybeans, or most other agricultural commodities, affords ample opportunity for maintaining speculative

* Emile Borel, "Valeur Pratique et Philosophie des Probabilités," *Traité du Calcul des Probabilités et de ses Applications*, Tome IV, Fascicule 3, ed. by Emile Borel (Gautier-Villars, Paris, 1950).

† Arrow, Karlin, and Scarf, *Studies in the Mathematical Theory of Inventory and Production* (Stanford, Calif.: Stanford University Press, 1958).

inventories. However, many companies automatically hedge their purchases in such fashion that they cannot achieve any speculative profit and, of course, cannot sustain any speculative loss. Under these circumstances the speculative motive is not existent and can be ignored in the analytical formulation of the inventory problem.

7. THE PLAN OF THE BOOK

Returning to the differentiations among inventory problems summarized at the end of Section 2, it is easy to calculate that there are 36 different classes, using only the five differentiations explicitly noted. Actually, there are other distinctions which are also of considerable consequence—a good example is the lengthy list of possible costs which may be involved—so the total number of subclasses of inventory problems is truly enormous. It is obvious that no one book could hope to deal with all of them. Nor would such a cataloging handbook of inventory models be particularly desirable. Even if a great number of models were included it would still frequently be the case that a specific, actual inventory problem would be somewhat different from any of the models in such a book. One of the fascinating features of inventory problems is their amazing diversity. Therefore, a more modest goal seems appropriate. We will attempt to present typical kinds of analyses of inventory problems so that the structure of the analytical formulation is made clear. We may then hope that the reader, faced with a different kind of problem, will be able to make the necessary modifications in the analysis himself.

Some kind of organization of approach is needed in order to lay bare the basic underlying unity of the different inventory models. Various alternatives are possible and, perhaps, there may be no "best" one. We have selected one which appears to us to be of considerable merit but we certainly do not claim that it is the best one. Essentially, we propose to make our first classification that of static and dynamic problems. Within this classification we will utilize the classification into certainty, risk, and uncertainty. Further classifications will be introduced when they seem desirable within this initial six class breakdown. Thus, we will begin with a presentation of static inventory models under certainty, risk, and, then, uncertainty. Parenthetically we may note that there is no problem in the first case. After this, we will proceed to a discussion of the dynamic inventory analyses with the same secondary breakdown.

This plan of attack refers to Part 1 of this book. Part 2 is devoted to the analysis of systemic factors in the resolution of inventory problems and the actual control of inventory situations. The use of computers in this regard has become so important in inventory problems that consideration must be given to the role which they can play. Part 2 will be devoted to an elucidation of this material including systems analysis, study design, data-processing and simulation.

static inventory problems under risk

8. GENERAL CHARACTERISTICS

The distinguishing characteristic of static inventory problems is that only one order is possible. This is certainly not the usual case in inventory problems but it is by no means rare. A dealer in Christmas trees, for example, may well have to place one initial order for all the trees he wants for the season. A newstand dealer may have to order as many racing forms as he wants for the day, with no opportunity for a second order. A retailer selling some seasonal item made, for example, in Japan may not be able to get a second order during the season because of the length of time involved in shipping. A retailer selling some kinds of high-fashion items may not be able to get a second order because of manufacturing problems. Some problems which are not, in strict fact, one-order cases are, nonetheless, effectively so because of the very large ordering cost. Such are, for example, inventory problems of plant capacity or schoolroom capacity, in the case of planning the size of a new school building. Such may also be the case in some kinds of lot-size problems. If, for example, a special kind of part has to be produced for a specific customer and to specified quality control standards and if the set-up cost is very large this may be the case. Because of the presence of defective items it is necessary to produce more items than the amount ordered. This, then, can be an inventory problem of the stated type.

Granting this defining characteristic, our chapter title would indicate that we have skipped our first class of problems: static inventory problems under certainty. The reason for this is that there is generally no analytical problem for this class of inventory situations. The amount needed is known and only one order is available. If it is profitable, or necessary, to have the given item then the known amount of it must be ordered. This statement needs some

slight qualification when due account is taken of the lag between order and delivery. There are two possibilities. A known total amount may be required, regardless of when it is delivered. This would be true, for example, for an inventory problem regarding some component needed for the construction of a building. If delivery were delayed then the construction would be delayed but the total amount would still be required. Alternatively, the known amount of demand might be on a per unit of time basis. In this case, it could happen that a time lag would require the ordering of a smaller amount because the demand during the time lag could not be met. If the amount of time lag is known with certainty all that is required is the subtraction of the indicated amount from the total order. If there is a probability distribution for the time lag we are really dealing with an example of the static inventory problem under risk. In short, there is no need for any treatment of the static inventory problem under certainty. In the case of the ordering of many items there may be, particularly in the case of the self-supplier, sequencing and scheduling problems of considerable intricacy. However, these kinds of problems are not inventory problems, they belong more properly to the subject of production scheduling and will not be treated here.

Most of the costs which we discussed previously can enter into the static inventory problem. However, there are two differences between static and dynamic problems with regard to costs which may be noted here. First, since there is only one order the cost of ordering is fixed for all courses of action except that of not ordering at all. In other words, the order cost is a fixed cost for all courses of action in such problems and, hence, can be ignored. Second, it is particularly in the case of the static problems that the question of salvage value—the overstock cost—arises. If any units of the item in question are left on hand at the conclusion of the demand period involved there will typically be some cost sustained in the disposal of them since the value will have deteriorated. This kind of cost can enter into dynamic inventory problems—specifically, those with a limited time horizon, like seasonal goods—but they must always be considered in the static inventory situation.

By definition, in this chapter we will be considering inventory problems where we know the probability distribution of demand. Generally, it is not a simple matter to determine such a probability distribution but we will not consider here any of the methods which might be used to accomplish this. Fundamentally, this is the forecasting problem and there is a huge literature dealing with the multitude of methods which have been developed in order to achieve successful forecasts. In Part 2 of this book we will have occasion to discuss some of the simpler methods which seem particularly apt for inventory problems. In the next chapter we will discuss more extensively some of the difficulties involved and suggest a possible compromise solution which will not require the detailed information about the probability distribution of demand which we are assuming here. For the purposes of this chapter we will be satisfied to discover the analysis which should be followed if we do have the necessary information about demand.

9. THE CHRISTMAS TREE PROBLEM

We will begin by considering a particularly straightforward kind of inventory problem which will permit us to make quite clear the analytical approach. We will try to do this by constructing what is called the payoff matrix for this inventory decision problem. We will consider the case of a merchant in a small town who wishes to stock some Christmas trees for sale during the Christmas season. We assume that he knows the probability distribution of demand for his Christmas trees during the season. Specifically, it is:

DEMAND	PROBABILITY
1	.05
2	.15
3	.20
4	.40
5	.10
6	.10
	1.00

The sum of the probabilities is, of course, unity since one of the indicated demands must occur. The meaning of the probabilities is as usual; e.g., the probability is .05 that the demand will be for one tree means that if the same situation were repeated for a great many seasons we could expect that in 5% of the seasons there would have been a demand for only one tree.

In addition, this merchant has some relevant cost and price information. He pays \$2 for each tree and he sells each tree for \$6. He delivers any tree which is sold and he pays \$.50 per tree to do this. Finally, if he has any trees left over at the end of the season he can sell them for firewood at \$.50 each. We will ignore any ordering cost. As we noted above, the ordering cost is fixed for all courses of action except for not ordering at all. Therefore, to ignore the ordering cost is equivalent to the procedure which would be followed if the merchant informed us that he had decided that he was definitely going to stock some Christmas trees and merely wanted us to determine how many he should order. We will also ignore carrying costs for the present. Finally, we will not consider loss of good will. Given, then, the demand probability distribution and the information concerning prices and costs, we want to determine the optimal number of Christmas trees for the merchant to order.

We may start by noting that the merchant has six reasonable courses of action, or strategies. He can order any number of trees from one to six. There is no possible reason for ordering more than six since he can never sell more than six and there is no possible reason for ordering less than one, ignoring the order cost, since he can always sell at least one. Furthermore, regardless of the number he orders there are six possible levels of demand: any number from one to six. In other words, if he orders only one tree there may still be a demand for any number from one to six. Of course, if he has only one tree he will only be able to meet the demand of one customer and

any other customers will have to buy their trees elsewhere. Since there are six alternative courses of action and six possible levels of demand it follows that there are a total of $6 \times 6 = 36$ possible combinations of one strategy and one level of demand. We want to begin by determining the merchant's payoff, the profit he makes, for each of these 36 possibilities. It is convenient to present these payoffs in the form of a payoff matrix: a rectangular array in which the strategies are represented by the rows, the possible levels of demand are the columns, and the payoffs are given in the cell at the intersection of the corresponding row and column. Such a payoff matrix will look like this:

		DEMAND					
		1	2	3	4	5	6
STRATEGY	ORDER						
	1	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
	2	F_{21}	F_{22}	F_{23}	F_{24}	F_{25}	F_{26}
	3	F_{31}	F_{32}	F_{33}	F_{34}	F_{35}	F_{36}
	4	F_{41}	F_{42}	F_{43}	F_{44}	F_{45}	F_{46}
	5	F_{51}	F_{52}	F_{53}	F_{54}	F_{55}	F_{56}
	6	F_{61}	F_{62}	F_{63}	F_{64}	F_{65}	F_{66}

The F 's stand for the payoffs: F_{32} , for example, represents the payoff which would result from the third strategy and from a demand of two. We want, of course, to determine the numerical values of these payoffs for our merchant's problem.

It is easy to calculate the payoff for any specific cell. Take the payoff which will result if the merchant orders one tree and there is a demand for one tree as a first example. He pays \$2 for the tree, sells it for \$6, and pays a delivery cost of \$.50. This leaves him with a profit (more strictly, a contribution to profit and overhead) of \$3.50. Thus, $F_{11} = \$3.50$. Suppose he orders four trees and there is a demand for only two. He pays \$8 for the four trees, he sells the two trees for \$12, he pays delivery charges of \$1, and he sells the two trees left over for \$1. This gives him a profit of \$4, which is F_{42} . We could continue calculating the individual payoffs in this manner but it is much quicker to systematize the procedure. Let us introduce two symbols:

x = amount ordered

z = demand level.

Using these symbols, let us make our calculations for three cases: demand less than, equal to, or greater than, the amount ordered. When demand is less than the amount ordered we will have the following contributions to the payoff: the merchant buys the trees for $2x$ dollars, he sells z of them for $6z$ dollars, he pays delivery charges of $.5z$ dollars, and he sells the leftovers for $.5(x - z)$ dollars. Adding these up we get a payoff of $(-1.5x + 5z)$ dollars. Using our symbols we would represent this case as that where $x > z$, where the sign is read "is greater than"—meaning that x , amount ordered, is greater than z , demand. When $x = z$ we have a simple calculation: he buys the trees for

$2x$, sells them for $6z = 6x$, and pays delivery charges of $.5z$. Adding this up gives a payoff of $3.5z$ or $3.5x$, as we prefer, since x and z are equal. The last case, where $x < z$, entails the following contributions to payoff: he pays $2x$ for the trees, he sells all of them for $6x$, and he has delivery charges of $.5x$. Thus, the total payoff will be $3.5x$. These calculations can be conveniently handled in a tabular form. We ascribe a positive sign to any inflow of cash and a negative sign to any outflow. Then the sum of all the relevant components will represent profit: profit in the usual sense if the sum is positive and a loss if the sum is negative. Finally, we note that we need consider only two different cases instead of three. This is so since the case where the amount ordered equals the demand ($x = z$) can equally well be incorporated into either of the other two cases: either becoming $x \geq z$ or $x \leq z$. So doing, we can summarize our calculations thus:

	$x \geq z$	$x < z$
Cost of trees	$-2x$	$-2x$
Sale of trees	$+6z$	$+6x$
Delivery of trees	$-.5z$	$-.5x$
Salvage of trees	$+.5(x - z)$	0
TOTAL	$-1.5x + 5z$	$+3.5x$

We can check these general expressions with the two payoffs we previously calculated. Thus, for $F11$ we have $x = z = 1$ and the first case applies so the payoff should be $-1.5(1) + 5(1) = \$3.50$, as before. For $F42$ we also have the first case, since $x = 4$ is greater than $z = 2$. Thus, the payoff should be $-1.5(4) + 5(2) = \$4$, also as before.

Using these general expressions it is easy to fill out the payoff matrix. We get:

ORDER	DEMAND					
	1	2	3	4	5	6
1	\$3.50	\$3.50	\$3.50	\$3.50	\$3.50	\$3.50
2	2.00	7.00	7.00	7.00	7.00	7.00
3	.50	5.50	10.50	10.50	10.50	10.50
4	-1.00	4.00	9.00	14.00	14.00	14.00
5	-2.50	2.50	7.50	12.50	17.50	17.50
6	-4.00	1.00	6.00	11.00	16.00	21.00

This table, then, shows the payoff which would result from any combination of a selection of a strategy and the occurrence of a specific level of demand. We will reserve for further comment the question of whether or not this table "looks right." For the moment we will merely note that our calculations of the payoffs are straightforward enough and that they have a clear economic meaning. In each case the payoff given is the amount of profit, or loss, which

will result if the merchant chooses that specific strategy and that specific level of demand occurs.

Now, of course, the whole essence of this kind of inventory problem is that we do not know what level of demand will, in fact, occur. We do, however, know the probability distribution of demand—the probability with which each level of demand can be expected to occur. We propose, then, to calculate, for each strategy, the expected value of the payoff for that strategy on the basis of the known probability distribution of demand. The expected value, it will be remembered, is simply the old-fashioned arithmetic mean of a set of values. The procedure for calculating the expected values is simple: for any given strategy we must multiply each possible payoff for that strategy by the corresponding probability of the given level of demand and add all of these products up. This is the expected value for the strategy. Thus, for our merchant's example we have for the first strategy:

$$.05(3.50) + .15(3.50) + .20(3.50) + .40(3.50) + .10(3.50) + .10(3.50) = 3.50$$

For the sixth strategy we would have:

$$.05(-4.00) + .15(1.00) + .20(6.00) + .40(11.00) + .10(16.00) + .10(21.00) = 9.25$$

Proceeding similarly we calculate all of the expected values:

ORDER	EXPECTED VALUE
1	\$3.50
2	6.75
3	9.25
4	10.75
5	10.25
6	9.25

We propose to advise the merchant to select that strategy which has the largest expected value of payoff, or just expected payoff for short. Why? We can present the justification for this approach quite succinctly. Suppose that this identical inventory decision problem were repeated a great many times and that each time the merchant had to reach a decision as to how many trees to order. It can be shown that if he selects, each time, the strategy with the largest possible expected payoff, then over the entire series of decisions he will make a larger total profit than by any other procedure. There are two immediate objections to this argument. The first argument says that it is obviously a fiction to imagine that this same decision problem will ever be presented to the merchant again—after all, nothing ever repeats itself *exactly*. Therefore, this argument continues, it is absurd to recommend that the merchant should govern himself on the basis of something that will never happen. The answer to this argument is that we can show that even if the merchant is presented with a series of different inventory decision problems he will make more profit in the long run if he chooses each time that

strategy with the largest expected payoff. The second argument is really a variant of the first. It rebuts our answer to the first argument by insisting that we shouldn't reason in terms of a series of inventory decision problems at all—each one of them is a unique and separate special case, to be coped with only on its own terms. This argument gets its venom from a continuation of it: suppose the merchant goes bankrupt on one specific inventory decision problem. What good, then, will be the hypothetical series of future decisions since he now has no more capital and is out of business? We hope the reader will already have recognized that we eliminated this argument by fiat in our discussion in the first chapter. The argument hinges around the fact that for some decision problems amounts of money are not suitable measures of payoff—some measure of utility is needed instead. We have maintained that this is rarely the case in the frequently occurring kinds of inventory problems. Therefore, the argument will not be valid against the use of expected payoffs in inventory decision problems. So we conclude that we are justified in recommending the choice of that strategy which has the largest expected payoff. In our example this is the strategy of ordering four trees, with an expected payoff of \$10.75. This, then, is the strategy the merchant should select. Obviously, he may not make as much as this or he may make more in the specific case. The outcome depends strictly on what level of demand eventuates. However, if the inventory problem were repeated many times the merchant could expect to average \$10.75 profit each season if he always ordered four trees.

Now that we have formulated, and resolved, this inventory decision problem, let us compare what we have done with our general comments concerning the approach to inventory problems. We talked previously of opposing costs, of the necessity of considering the total costs which are equal to the sum of the various opposing costs, and of the minimization of the total costs. When we review our present example it appears that the only correspondence there is with this suggested scheme is that we did, in fact, minimize something. However, there is no immediate indication that we minimized anything called total costs and, further, there is no sign of any pair of opposing costs. This seems passing strange if there is any merit in our previous generalizations.

Finally, if we look at the payoff matrix itself—with an eye not too jaundiced by knowledge of economic principles—we may observe that it appears to be a questionable representation of this inventory decision problem. While we have explicitly disclaimed any cost due to the loss of customer goodwill there still seems to be something wrong with the matrix. We are referring to the fact that we have never ascribed a cost to reflect sales which we might have had if we had ordered more stock. Thus, according to our payoff matrix there is a payoff of \$3.50 for the first strategy no matter what the demand level. Yet if the demand were for six trees we could have made a payoff of \$21.00, if we had had six trees in stock. Surely there should be some cost associated with such a loss of possible profit. We must consider these comments on our resolution of this inventory problem in some greater detail in the next section.

10. THE OPPORTUNITY COST MATRIX

The answers to the objections in both the two preceding paragraphs stem from the same fundamental conception—opportunity costs. In summary, the answer is that we are, in fact, ascribing a cost to the loss of possible profit and that when this cost is taken into account we are, in fact, balancing opposing costs or, what is equivalent, minimizing total costs. We must now try to demonstrate that this is the case.

Let us consider the question of a cost due to a profit that might have been made if there had been more stock ordered. What this amounts to is that after the actual demand has occurred it may well turn out that we could have made more profit if we had originally selected some strategy other than the one we did select. In short, we lost an opportunity—hence the name, opportunity costs. Suppose, for example, that we ordered six trees and the demand was only for one tree. Our payoff matrix shows that the payoff in this case is $-\$4$, a loss of $\$4$. However, we could have made a profit of $\$3.50$ if we had only ordered one tree and the demand had been for one tree. Therefore, we can reason that the true cost of our decision to stock six trees—in the event of a demand of one tree—would be $\$7.50$, equal to a $\$4$ loss actually sustained plus the $\$3.50$ profit we might have made. Suppose we ordered one tree and the demand was for six trees. By a similar argument we should credit our decision with a cost of $\$17.50$, equal to the $\$21.00$ we would have made if we had ordered six trees minus the $\$3.50$ we actually made as a result of ordering only one tree. These are the opportunity costs for these two cases. An opportunity cost is, in short, a cost sustained because the decision taken was not the best in terms of the level of demand which actually occurs. The opportunity cost of the best decision for the given level of demand is, then, by definition equal to zero. Our previous payoff matrix happened to be expressed in terms of profits but we shall find that it is often the case that the payoff matrix may be most directly expressed in terms of costs. In either case we can calculate opportunity costs. If the original payoffs are in terms of profits as in our example, we simply subtract any specific payoff from the highest payoff in the same column in order to get the opportunity cost. If the payoffs are in terms of costs we need only subtract from any specific payoff the lowest payoff in the same column of the payoff matrix in order to get the opportunity cost. Since there is an opportunity cost for each combination of a strategy and a level of demand we can present all the opportunity costs in the form of a payoff matrix also. It is easy to calculate the opportunity cost matrix directly from the original payoff matrix. Take any column of the payoff matrix—corresponding to a specific level of demand—and find the largest payoff if the payoffs are profits, the smallest payoff if the payoffs are costs. Then subtract each payoff in the same column from the largest payoff to get the corresponding opportunity costs in the case of profits. For costs, subtract the smallest payoff from each payoff in the same column to get the opportunity costs. In our example we get the opportunity cost matrix:

ORDER	DEMAND					
	1	2	3	4	5	6
1	\$0	\$3.50	\$7.00	\$10.50	\$14.00	\$17.50
2	1.50	0	3.50	7.00	10.50	14.00
3	3.00	1.50	0	3.50	7.00	10.50
4	4.50	3.00	1.50	0	3.50	7.00
5	6.00	4.50	3.00	1.50	0	3.50
6	7.50	6.00	4.50	3.00	1.50	0

It can be seen that in this opportunity cost matrix the best decision for each level of demand is credited with a cost of zero. Any other decision is ascribed a cost because of the lost opportunity involved due to having missed the best choice. In particular, every case where the ordered quantity is less than the actual demand sustains a cost of \$3.50 for each tree too few. This is precisely what our objection at the end of the preceding section demanded.

This opportunity cost matrix seems like a perfectly satisfactory way to represent the merchant's inventory problem. As a matter of fact, the ascription of costs in the opportunity cost matrix may seem more reasonable than the original payoff matrix. Let us, then, use the opportunity cost matrix as the basis for our analysis of the inventory problem. Once again, we don't know which specific level of demand will occur but we do know the probability distribution of demand. Therefore, we can use exactly the same argument as before to justify the calculation, and use, of expected values. In this case we will be calculating expected opportunity costs. The only difference will be that since these are costs the merchant will want to select that strategy which will minimize his expected opportunity cost rather than maximizing as he did for expected profits. The calculations of expected values proceed exactly as before. For the first strategy we have:

$$.05(0) + .15(3.50) + .20(7.00) + .40(10.50) + .10(14.00) + .10(17.50) = \$9.275$$

For the sixth strategy we have:

$$.05(7.50) + .15(6.00) + .20(4.50) + .40(3.00) + .10(1.50) + .10(0) = \$3.525$$

Proceeding similarly we calculate all of the expected values:

ORDER	EXPECTED OPPORTUNITY COST
1	\$9.275
2	6.025
3	3.525
4	2.025
5	2.525
6	3.525

In order to minimize his expected opportunity costs the merchant must select his strategy of ordering four trees. We remember immediately that this was exactly the same conclusion—order four trees—we reached when we analyzed the original payoff matrix. Is this a coincidence?

Of course, it is not a coincidence. The analysis by means of opportunity costs will always produce exactly the same conclusion as will the analysis in terms of total costs or total profits as payoffs.* The expected values calculated in the two cases will certainly be different but the strategy which will be selected will be the same in both cases. Let us see why this is so. We derived the individual opportunity costs directly from the original payoff matrix. Therefore, we can write the calculation for a given strategy's expected opportunity cost directly in terms of the entries in the original payoff matrix. To do this all we need do is to replace each opportunity cost by the corresponding calculation we performed on the original payoffs in order to get the opportunity cost. Thus, consider the opportunity costs for the first strategy. We know that:

$$0 = 3.50 - 3.50$$

$$3.50 = 7.00 - 3.50$$

$$7.00 = 10.50 - 3.50$$

and so forth. Therefore, our calculation of the expected opportunity cost for this strategy could have been written:

$$\begin{aligned} &.05(3.50 - 3.50) + .15(7.00 - 3.50) + .20(10.50 - 3.50) \\ &+ .40(14.00 - 3.50) + .10(17.50 - 3.50) + .10(21.00 - 3.50) = 9.275 \end{aligned}$$

as before. But we could equally well rearrange these terms and write the left-hand side of the above equation as:

$$\begin{aligned} &[.05(3.50) + .15(7.00) + .20(10.50) + .40(14.00) + .10(17.50) + .10(21.00)] \\ &- [.05(3.50) + .15(3.50) + .20(3.50) + .40(3.50) + .10(3.50) + .10(3.50)] \\ &= 12.775 - 3.50 = 9.275 \end{aligned}$$

We can write the expected opportunity cost for each strategy in this form. Inspection of the procedure is sufficient to show that, if we did so, the whole positive term in the preceding equation will occur in each expected opportunity cost equation in our example. The negative term for each expected opportunity cost, written in the preceding form, is nothing more nor less than the expected payoff for the given strategy in the original payoff matrix. In other words, for our example, we have that:

$$\begin{aligned} &\text{Expected opportunity cost for a given strategy} \\ &= \$12.775 - \text{Expected payoff for that strategy} \end{aligned}$$

Comparison of the original expected payoffs with the expected opportunity costs will show that this relationship is indeed true. In general, for any payoff matrix in terms of profits we have:

$$\text{E.O.C.} = K - \text{E.F.}$$

where E.O.C. = expected opportunity cost; K = a constant; and E.F. = expected payoff (profit).

* See page 54.

Granted this relationship, we can see why the two approaches will always produce the same selection of strategy. In the original payoff matrix we found that strategy which maximized E.F. It is obvious from the equation above that the maximum value of E.F. will simultaneously produce the minimum value of E.O.C. and, of course, when we use the opportunity cost matrix that is precisely what we do—select the strategy which has the minimum E.O.C. It is clear, then, that the two analyses will always produce the same conclusion. If the original payoff matrix is in terms of costs it can be shown, by exactly similar reasoning, that the relationship will be of the form:

$$\text{E.O.C.} = \text{E.F.} - K$$

where E.F. is now in terms of costs. In this kind of case the procedure would be to select that strategy, in the original payoff matrix, which had the minimum cost. It can be seen from the preceding equation that the same strategy would simultaneously have the minimum E.O.C. so, once again, the two approaches would produce the same conclusion.

This analysis shows that our second criticism of the original payoff matrix analysis—that it took no account of lost potential profits—was not correct. The original payoff matrix does take account of lost potential profits, but in a somewhat disguised form. Further, once we have converted to the opportunity cost matrix we can see immediately that two opposing costs are, in fact, involved and that we are minimizing the sum of these costs. Or, more precisely, we are minimizing the expected value of the total costs, which is the best we can do when we are dealing with inventory problems under risk. The opposing costs are evident: a cost of \$1.50 per tree overstocked against a lost profit, or cost, of \$3.50 for each tree understocked. These costs are opposing since the more trees we stock the less the risk of sustaining the lost profit cost but the greater the risk of sustaining the overstock cost. Thus, our first objection to the original analysis is also mistaken. The general approach to inventory problem analysis which we previously outlined is, in fact, being followed. That this is the case is not obvious in the payoff matrix analysis but is highlighted by the change to the equivalent opportunity cost matrix.

As a matter of fact, the explicit presentation of the two opposing costs in the form of opportunity costs permits us to analyze this kind of inventory problem in a much simpler way called incremental analysis. It is simpler because it avoids the calculation of the E.F.'s or the E.O.C.'s as we did above. Let us conceptualize the decision problem a bit differently. Suppose the merchant made not one, but a series of decisions in this self-same problem. Specifically, suppose that he were first to choose between ordering one tree or not ordering it. If he decides to order one he then decides whether he will order one more or not. He proceeds similarly until he reaches the point where he decides not to order the next tree. This ends his series of decisions and he orders the immediately preceding number of trees. How might this kind of a decision process be analyzed? In order to consider this we need to introduce some notations dealing with probabilities. We will let $P()$

designate the probability of whatever is written in the brackets. Thus, using our demand distribution:

$$P(z = 3) = .20$$

is to be read "the probability that demand, z , equals 3 is .20."

$$P(z \geq 3) = .80$$

is to be read "the probability that demand, z , is greater than or equal to 3 is .80."

$$P(z < 3) = .20$$

is to be read "the probability that demand, z , is less than 3 is .20."

Now, how can we present the opportunity cost matrix for the decision problem as just described? Suppose the merchant has reached the i th decision: he is considering whether to order the i th tree or not. He has only two possible strategies: to order it or not to order it. And there are only two relevant possibilities regarding demand: either there will be a demand for the i th tree or there will not be. We can, therefore, present this decision problem in a two-by-two opportunity cost matrix.

	DEMAND FOR i TH TREE	NO DEMAND FOR i TH TREE
Order i th tree	\$0	\$1.50
Don't order i th tree	3.50	0

The opportunity costs entered in the matrix were already determined but we can calculate them directly. If he orders the i th tree and there is a demand for it, there is clearly an opportunity cost of zero because he has made the best of his two possible decisions. If he orders the i th tree and there is no demand for it he has not made the best decision and he sustains a cost of \$1.50, the loss on an overstocked tree. Similarly, if there is demand for the i th tree and he did not order it he has not made the best decision and he sustains an opportunity cost of \$3.50, the profit he might have had if he had ordered it. And, of course, the fourth entry in the matrix is zero by a similar argument.

How can we analyze this opportunity cost matrix? Exactly as before—by calculating the E.O.C. for each possible strategy. There will be a demand for the i th tree if the demand level is i or greater. Using our notation, this has probability of $P(z \geq i)$. Since there either is demand for the i th tree or there isn't, and since the sum of all probabilities must always equal one, it follows that the probability that there will not be a demand for the i th tree is simply $[1 - P(z \geq i)]$. Thus, the two E.O.C.'s are simply:

Order: $(0) P(z \geq i) + (1.50)[1 - P(z \geq i)] = 1.50 - 1.50 P(z \geq i)$

Don't order: $(3.50) P(z \geq i) + (0)[1 - P(z \geq i)] = 3.50 P(z \geq i)$

Now, the merchant will order if, and only if, the E.O.C. for ordering is less

than the E.O.C. for not ordering. Algebraically, this means that he will order if and only if

$$1.50 - 1.50 P(z \geq i) < 3.50 P(z \geq i)$$

or

$$1.50 < 5.00 P(z \geq i)$$

or

$$.30 < P(z \geq i)$$

In words, he will order the i th item only if $P(z \geq i)$ is greater than .30. This can be immediately determined from the probability distribution of demand. All that is necessary is to make a new column for the table showing this distribution. The new column shows $P(z \geq i)$ for each possible value of i . Thus:

DEMAND		
i	$P(z = i)$	$P(z \geq i)$
1	.05	1.00
2	.15	.95
3	.20	.80
4	.40	.60
5	.10	.20
6	.10	.10

We see that $P(z \geq 4) = .60$ is greater than .30 but $P(z \geq 5) = .20$ is not greater than .30. Therefore, in accordance with our analysis we would order the fourth tree but not the fifth. Thus we reach exactly the same conclusion as before but by a route requiring considerably less arithmetic. One merit of this last analysis is that it makes the opposing costs and their relation to the analysis quite apparent.

It is clear that we can equally well analyze a given inventory problem either in terms of a payoff matrix based on total profits or total costs or else in terms of an opportunity cost matrix. The two forms of analysis are equivalent and will always produce the same choice of a strategy. It is essential, however, that the distinction between the two approaches should be kept clearly in mind. It is fatal to confuse the two so that we end up with a supposed payoff matrix which has some entries in terms of total profits or costs and other entries in terms of opportunity costs. We must be consistent in our approach. That it is easy to confuse the two is indicated by the fact that there are published examples where precisely this has been done.

11. THE COST OF RISK

In our previous discussion we showed that we always have the following relationship between expected opportunity cost and expected profit;

$$\text{E.O.C.} = K - \text{E.F.}$$

Let us give a little further consideration to the K in this equation. In our example we had $K = \$12.775$. More specifically, K was calculated by taking

the expected value of the largest entries in each column. In other words, K is the expected value of the payoffs we would receive if we always selected the best course of action.

This can be stated more picturesquely as follows. Suppose that we were omniscient—at least in regard to the inventory problem under consideration. We know, in short, exactly what level of demand will occur and we know it in advance so that we can select our strategy accordingly. If this were the case we would certainly always select that strategy which would give us the largest payoff. Now, this kind of hypothetical case shouldn't be confused with the situation in which we were also omnipotent. If we were omnipotent we could control which level of demand would occur and, of course, we would always arrange things so that there would be a demand for six trees. We would always stock six trees and, assuming the same decision problem were repeated, we would always make a profit of \$21.00. We are not assuming so much. Being omniscient gives us no control over the actual level of demand—we simply know what it will be. Thus, sometimes we will make \$21.00—if the demand happens to be for six trees—but sometimes we will make only \$17.50, \$14.00, and so forth. We will always select the best strategy so we will always make the maximum amount for the given level of demand.

Under these circumstances, what would be our expected profit over a series of identical decision problems? The amount we would make on any specific decision would depend on the level of demand and we know the probabilities of the various levels of demand. Therefore, to determine our expected profit all we need do is to calculate the expected value of the amounts we would make for each level of demand:

$$.05(3.50) + .15(7.00) + .20(10.50) + .40(14.00) + .10(17.50) + .10(21.00) = 12.775 = K$$

Thus, K is simply the average amount we would make if we were omniscient or, a little more prosaically, K is the average amount we would make if we had perfect information.

Returning to our original calculations of the expected profits for our example we see that the largest expected profit was \$10.75, for the strategy of ordering four trees. We have just seen that we could make an expected profit of \$12.775 if we had perfect information but with the information we do have we can make an expected profit of \$10.75. The difference between these two amounts is $12.775 - 10.75 = 2.025$. This, then, measures the upper limit of the worth to the decision-maker of further information about the demand level. It will be noted, of course, that this amount, \$2.025, is the E.O.C. for the strategy of ordering four trees. Our argument is perfectly general—with obvious changes for the case where the original payoffs are costs—and gives us another interpretation of an E.O.C. The E.O.C. of the optimal strategy is the upper limit of the worth to the decision-maker of further information about the demand level. Put in the other direction we can say that the E.O.C. of the optimal strategy is the cost to the decision-maker of his lack of perfect information about the level of demand. This suggests that it might be possible to evaluate the potential worth of market surveys or

demand analyses undertaken to gain a better knowledge of future demand level. As a matter of fact, this is so. Argumentation along these lines has been developed—and in some very ingenious ways—by Robert Schlaifer in his *Probability and Statistics for Business Decisions*.^{*} We refer the reader to that book for a development of some of the implications of this approach.

12. OTHER COST STRUCTURES

The analysis of this kind of inventory situation can equally easily be extended to include other kinds of costs which may be relevant. There is no particular difficulty in so doing except that there may be questions about how to apply some of the costs, particularly the carrying cost. In order to illustrate the procedure—even though it is somewhat repetitious—let us take an example with which we can also indicate how payoffs and opportunity costs can sometimes be confused. Consider the inventory problem of a manufacturer who purchases an “orphan” machine tool—one of which the maker is going out of business. Let us suppose that there is a particular part of this machine tool which is subject to failure. When the manufacturer purchases the tool he can buy spares for this part at a cost of \$100 each. If there is a failure of this part and the manufacturer does not have a spare available then one will have to be specially machined at a total cost, including the resulting down time of the machine tool, of \$1000. The manufacturer estimates the lifetime of the tool will be ten years and from his experience with similar tools he has the following probability distribution of number of failures of this part during the life of the tool:

FAILURES	PROBABILITY
0	.30
1	.40
2	.20
3	.10
	1.00

This probability distribution represents, of course, the demand for the spare parts in question. We will ignore carrying costs for the moment and we will assume that unused parts have no salvage value. This means that the major difference between this problem and our Christmas tree example is that here we are reasoning in terms of costs, which the manufacturer naturally wants to minimize, instead of maximizing profits.

Let us construct a matrix for this problem directly along the lines of opportunity costs. We need consider only four strategies—ordering from zero to three spare parts. For each strategy we will assign a zero cost if exactly that number of parts is, in fact, needed. If we have fewer parts than we need it will cost us an additional \$1000 per part that we are short. If we have

^{*} Robert Schlaifer, *Probability and Statistics for Business Decisions* (New York: McGraw-Hill Book Co., Inc., 1959).

more parts than we need it will cost us the wasted \$100 per extra part. This simple analysis immediately gives the following matrix:

ORDER	FAILURES			
	0	1	2	3
0	\$0	\$1000	\$2000	\$3000
1	100	0	1000	2000
2	200	100	0	1000
3	300	200	100	0

Now, the only trouble with this matrix, which looks quite reasonable, is that it is wrong! We have committed precisely the error we warned against—confusing total costs and opportunity costs in one matrix. It is a very easy mistake to make! Let us see if we can discover exactly how we happened to blunder into this error. We did it by surreptitiously changing the idea of opportunity cost to additional cost. Now, it is true that opportunity costs are additional costs but it all depends on what we measure them against—what are they additional to? In the present case we have changed our base in measuring the additional costs and this is why we have blundered. Specifically, when failures are greater than the amount ordered we have measured our costs as additional to what has already been spent. These are not opportunity costs. When failures are less than the number of spares ordered we have measured our costs as additional to what they might have been. These are opportunity costs. We simply haven't been consistent for the entire set of costs.

It is, of course, a simple matter to get the correct matrix, either payoff in costs or opportunity costs. First, let us determine the total cost matrix. Take the strategy of ordering one spare as an example. We pay \$100 for the spare and that is all we pay unless there are two or three failures. If there are two failures we need one extra spare at a cost of \$1000 or a total cost of \$1100. For three failures the total cost will be, \$2100. Proceeding similarly, we can immediately fill in the total cost matrix:

ORDER	FAILURES			
	0	1	2	3
0	\$0	\$1000	\$2000	\$3000
1	100	100	1100	2100
2	200	200	200	1200
3	300	300	300	300

This is the correct total cost matrix. The opportunity cost matrix can be calculated directly from the total cost matrix:

ORDER	FAILURES			
	0	1	2	3
0	\$0	\$900	\$1800	\$2700
1	100	0	900	1800
2	200	100	0	900
3	300	200	100	0

We will leave to the reader the determination of the optimal strategy. This illustrates the way in which the opportunity costs can be plausibly confused. We will now add some carrying costs to our original formulation of this example to see how they should be included in the matrix.

Suppose that the manufacturer in question considers that his carrying cost—due primarily to the loss of return on the capital tied up in the inventory—is 10% per year. How can we incorporate this cost in our representation of his inventory problem? There is no difficulty in calculating this cost for the spares which we buy and never use: \$100 at 10% for 10 years is very nearly \$259. This is the cost of an unused spare. Or is it? What about discounting future sums of money to present worth? And how is \$259 in any sense the cost of a spare part which cost only \$100? Let us proceed more slowly. The \$259 is obviously only a cost in the sense of an opportunity cost. This suggests that we should consider that the manufacturer has a certain sum of money which he can either spend on spares or which he can alternatively utilize. In the latter case he would have, starting with \$100, \$259 at the end of ten years while if he invests the \$100 in a spare which will not be used he will have zero at the end of ten years. So we can equally well calculate the cost as \$100 now or as \$259 at the end of ten years. We must, in short, do all of our calculations on the basis of now or on the basis of ten years hence. We will elect to make the calculations on the basis of present worth (the “now” alternative).

In order to determine the carrying costs involved when less spares are ordered than are needed we must make some assumption about the times when failures occur. The most natural such assumption is that they are equally spaced in time. Thus, if one failure occurs during the machine tool's lifetime we will assume that it occurred after five years, if two failures occurred we will assume they occurred after $3\frac{1}{2}$ and $6\frac{1}{2}$ years, and so forth. Suppose we don't order any spares and there is one failure. It will cost us \$1000 five years hence and, at a discount rate of 10%, this has a present worth of \$621 [$= 1000/(1.10)^5$]. If there were two failures it would cost \$1000 in $3\frac{1}{2}$ years, a present worth of \$728, and \$1000 in $6\frac{1}{2}$ years, a present worth of \$530. This is a total present cost of \$1258. Similarly, if there were three failures the present worth of the total cost would be \$1898. In this fashion, at the cost of some arithmetic, we complete the following total cost matrix—where all costs are at present worth:

ORDER	FAILURES			
	0	1	2	3
0	\$0	\$621	\$1258	\$1898
1	100	100	630	1210
2	200	200	200	689
3	300	300	300	300

We can now use the probability distribution of failures to calculate the expected value for each strategy. We find:

ORDER	EXPECTED VALUE
0	\$689.80
1	317.00
2	248.90
3	300.00

Since the payoffs are in terms of costs the decision-maker wants to minimize the expected value. Therefore, he would select the strategy of ordering two spare parts.

13. THE MATHEMATICAL FORMULATION: THE DISCRETE CASE

All of our analysis so far has been handled arithmetically. This has the advantage of making the logical structure of the analysis quite clear but it has some disadvantages. Our examples have been ones where there were only a few possible levels of demand. More typical are kinds of inventory problems where there are a great many possible levels of demand—far more than we would like to present explicitly in a payoff matrix. We might be able to avoid the payoff matrix representation of a large inventory problem by using the incremental analysis we developed above. However, even this is sometimes not very convenient because we may be presented with the demand probability distribution in the form of one of the well-known basic kinds of probability distributions: normal, binomial, Poisson, and so forth. Further, some kinds of inventory problems do not involve stock in units. For example, inventory problems involving liquids are basically continuous, as compared to the discrete kinds of problems we have dealt with so far. For these reasons it is desirable to have a general formulation of this kind of inventory problem. This will require the introduction of some mathematical notation but the logic of the analysis remains the same.

For this relatively simple kind of inventory problem—a static problem under risk—the major advantages of a mathematical formulation only result when we introduce continuous types of probability distributions, either because the inventory has a continuous nature or as an approximation to the discrete kind of inventory problem. However, it may be useful to indicate how the discrete case can be mathematically formulated. This will provide a direct parallel to the arithmetical calculations we have been using and will be an intermediate step towards the use of calculus in the continuous cases. Let us, then, reconsider the example of the preceding section. We will consider this example without any carrying costs. Under these circumstances there are two opposing costs: the cost of stocking and the cost of an understock. In our example these costs were \$100 and \$1000, respectively. Note that this is in the sense of total costs, not opportunity costs. Let us generalize to the degree that we have:

c = total cost of one unit

C_u = total cost of an understock of one unit

For our formulation we need to use a mathematical symbol of the kind called

operators. Such symbols do not function like the usual x 's and y 's of algebra. They are shorthand instructions to do something to, to "operate on," the more usual mathematical variables which follow them in the equation. The one we need is $\sum$, read "sigma," which designates summation. Its presence in an equation means that whatever comes after it must be added up. We can illustrate its use on some symbols we have already used. For greater ease of writing, and reading, in what follows we will replace our previous $P(z = j)$, the probability that demand, z , equals the specific value j , by P_j . This shorter form is equally unambiguous and simplifies the appearance of our equations. Now, by using our new operator we can express some of the other probabilities we introduced quite simply:

$$P(z \leq j) = \sum_{i=0}^j P_i = 1 - \sum_{i=j+1}^N P_i$$

$$P(z \geq j) = \sum_{i=j}^N P_i = 1 - \sum_{i=0}^{j-1} P_i$$

where N is the largest possible demand. The operator only tells us to add so the entries below and above the operator are necessary to give us the further information of how many things to add. The use of i is merely to give this information and thus i has no intrinsic mathematical significance in the equation. For this reason i is called a dummy of summation. The first equation, translated into words, tells us that to calculate the probability that z is less than or equal to j we must add up the probabilities that z equals zero, that z equals one, and so forth, up to and including the probability that z equals j . The second half of the equation utilizes the fact that the sum of all the probabilities must be one in order to give a different, but equivalent, way of calculating the same probability.

Our procedure was to calculate the expected values for each strategy. Letting $E.F_k$ designate the expected payoff for the k th strategy we have:

$$E.F_k = (F_{k0})P_0 + (F_{k1})P_1 + \dots + (F_{kN})P_N$$

where N is the largest possible demand level, or, using our summation operator:

$$E.F_k = \sum_{i=0}^N F_{ki}P_i$$

We can introduce, for our specific example, the algebraic form of the payoff. This takes two forms, depending on whether the amount ordered, k , was less than or greater than the demand level, j . We have:

$$\begin{aligned} F_{kj} &= kc && \text{if } k \geq j \\ &= kc + (j - k)C_u && \text{if } k < j \end{aligned}$$

Substituting these expressions in the equation for $E.F_k$ gives:

$$E.F_k = \sum_{i=0}^k kcP_i + \sum_{i=k+1}^N [kc + (i - k)C_u]P_i$$

$$\begin{aligned}
 &= kc \sum_{i=0}^k P_i + kc \sum_{i=k+1}^N P_i + C_u \sum_{i=k+1}^N iP_i - kC_u \sum_{i=k+1}^N P_i \\
 &= kc + C_u \sum_{i=k+1}^N iP_i - kC_u \sum_{i=k+1}^N P_i
 \end{aligned}$$

This equation represents exactly the same thing as we calculated for the expected total cost payoff for the example in question. Previously we simply calculated all of the expected values and selected that one which was the minimum. Can we do the same thing mathematically? Yes, we can but the calculations are not nearly so neat as they are when we are dealing with continuous probability distributions and can therefore use the methods of calculus. We reason as follows. We want to find the specific strategy which has the smallest expected total cost. This strategy is designated by some value of k , let us call it K . If this strategy has the smallest expected total cost it follows, in particular, that the two adjoining strategies, $(K+1)$ and $(K-1)$, have larger total costs. The expected total costs for these two strategies are:

$$E.F_{K+1} = (K+1)c + C_u \sum_{i=K+2}^N iP_i - (K+1)C_u \sum_{i=K+2}^N P_i$$

and

$$E.F_{K-1} = (K-1)c + C_u \sum_{i=K}^N iP_i - (K-1)C_u \sum_{i=K}^N P_i$$

We can write these two expected total costs in terms of the expected total cost of K , thus:

$$E.F_{K+1} = E.F_K + c - C_u \sum_{i=K+1}^N P_i$$

and

$$E.F_{K-1} = E.F_K - c + C_u \sum_{i=K}^N P_i$$

It can be seen from the form of these relationships that any strategy of ordering more than $K+1$ will have a still larger expected total cost than that of ordering $K+1$ since the positive term in the expression remains fixed while the negative term gets smaller as larger amounts are ordered. Similarly, ordering any amount less than $K-1$ will produce a still larger expected total cost than that of ordering $K-1$ since the negative term in the expression remains fixed while the positive term gets larger. Therefore, if the expected total cost of the strategy of ordering K is, in fact, smaller than the expected total costs of ordering $K+1$ and of ordering $K-1$, then it will be smaller than any other strategy's expected total cost. We want to find the strategy, K , for which this is true. It will be true if the following holds:

$$E.F_{K+1} - E.F_K = c - C_u \sum_{i=K+1}^N P_i > 0$$

and

$$E.F_{K-1} - E.F_K = -c + C_u \sum_{i=K}^N P_i > 0$$

These will be simultaneously true if and only if

$$\sum_{i=K}^N P_i > \frac{c}{C_u} > \sum_{i=K+1}^N P_i$$

Using our former notation, this is equivalent to

$$P(z \geq K) > \frac{c}{C_u} > P(z \geq K + 1)$$

In order to use this relationship we need to know the probabilities of the form $P(z \geq k)$. As we did before, we can determine these directly from the probability distribution of demand. For our example we have:

DEMAND		
k	$P(z = k)$	$P(z \geq k)$
0	.3	1.0
1	.4	.7
2	.2	.3
3	.1	.1

For our example, $c/C_u = \$100/\$1000 = .1$. Our relationship is expressed in terms of inequalities but in the present case we see that an equality holds. In other words, we have either:

$$P(z \geq 2) > .1 = P(z \geq 3)$$

or

$$P(z \geq 3) = .1 > P(z \geq 4)$$

This causes no difficulty. It means that there are two strategies with the same expected total cost and we can equally well use either one of them. In the present case we can either order two or three spares. In both cases, as the reader can verify, the expected total cost is \$300.

It will be noted that there is a considerable degree of similarity between the conclusion of this analysis in terms of total costs and the earlier incremental analysis we developed in terms of opportunity costs. At least they take a similar form in expressing the conditions that must be met by the optimal strategy. We developed the incremental analysis for the Christmas tree problem, for which the original payoff matrix was expressed in terms of profits. Our present analysis was developed in terms of a total cost matrix and, hence, is different from one developed for a profit matrix. Therefore, we can make no direct comparison of the two analyses on the Christmas tree problem. However, we can use the incremental opportunity cost analysis on our present spare-part example. For the spare-part problem the opportunity cost of an overstock is \$100 and the opportunity cost of an understock is \$900. The incremental analysis therefore shows that we should select the last strategy for which $P(z \geq j)$ is greater than $\$100/(\$900 + \$100) = .1$. Equality instead of inequality for the incremental analysis has the same effect as was just discussed above for the analysis of this section. Therefore, we see that the two analyses lead to exactly the same conclusion. Of course,

this is to be expected, if the analyses are valid. This short calculation simply demonstrates that it is so in the present case.

14. THE MATHEMATICAL FORMULATION: THE CONTINUOUS CASE

Our argument of the preceding section really consisted of two parts. First, we had to represent in the form of a mathematical expression the payoffs and the procedure for determining the expected payoff for a specific strategy. Second, we had to find some way to determine mathematically the minimum such expected payoff. When we turn to the representation of inventory problems by means of continuous probability distributions we find that the first part of the procedure is somewhat more unfamiliar, at least to those who are not quite conversant with calculus, but that the second part becomes a bit easier. Perhaps this in itself is another indication of the ubiquity of opposing costs!

The major new idea we need is that of continuous probability distributions. The probability distributions of demand which we have considered so far have been discrete ones. The probability of demand being, say, for four units is given and similarly with other discrete levels of demand. There is, however, no probability—nor, indeed, even possibility—that the demand would be for, say, $3\frac{1}{2}$ units or for 2.761 units. As we have already noted, there are two main reasons why such discrete distributions may not be satisfactory. First, the inventory in question may be for liquids or for lengths of copper wire or anything similar for which there may be a demand for any arbitrary amount. Second, for complex discrete inventory problems it is sometimes desirable, and sometimes almost necessary, to approximate the actual discrete probability distribution of demand by a continuous one in order to analytically resolve the problem. A continuous probability distribution of demand gives the probability that demand will equal any level, integral or fractional. We describe this mathematically by defining the probability that demand will fall between two levels separated by an “infinitesimal” distance. Thus, we define:

$$P(y < z < y + dy) = f(y) dy$$

where y is the dummy variable designating the level of demand, analogous to i in the discrete analysis, dy is the “infinitesimal” distance, z is, as usual, the actual level of demand, and $f(y)$ is some mathematical function which meets certain requirements enabling it to act as the probability distribution of demand. With the basic probabilities of demand so defined the analogy to the summation we used for the discrete case is integration, with its own operator sign, $\int$. Specifically, we define:

$$P(z \leq k) = \int_0^k f(y) dy = F(k)$$

and
$$P(z > k) = \int_k^{\infty} f(y) dy = 1 - \int_0^k f(y) dy = 1 - F(k)$$

where k is an arbitrary constant. The entries at the top and bottom of the integral sign are the limits of integration and fulfill the function of telling us how many probabilities have to be "added up," or integrated. The equivalence of the two forms:

$$\int_k^\infty f(y) dy = 1 - \int_0^k f(y) dy$$

follows from the fact that

$$\int_0^\infty f(y) dy = 1$$

is one of the major requirements that $f(y)$ must fulfill. This, of course, must be true for any probability distribution whatsoever; the sum, or integral, of all the probabilities must equal one.

Using these ideas the continuous analogue of the discrete case presented in the preceding section can easily be formulated. Two costs are involved: c = the total cost of one unit and C_u = the total cost of an understock of one unit. We are now assuming that fractional parts of a unit can be involved in stocks and understocks and that the corresponding proportion of the unit costs will measure the relevant costs. If we order x units it will cost us cx and this is all it will cost us unless there is an understock. In case of an understock there will be the original cost of cx plus the understock cost of $C_u(z - x)$, where z is the actual level of demand. To calculate the expected value of the cost for any given x we need to integrate the understock cost over all possible levels of demand which would produce an understock, each multiplied by the probability of that level of demand. This is accomplished thus:

$$E.F_x = cx + C_u \int_x^\infty (y - x)f(y) dy$$

We want to find the minimum such expected payoff (total cost) in terms of x . In other words, what value of x will give the minimum expected total cost?

To determine minima in calculus requires taking the derivative of the given expression and setting it equal to zero. In the present case, and in most of the continuous models of this kind of inventory problem, the expected payoff will include integrals. This is the only somewhat unusual feature of this kind of minimization problem. How do we take derivatives of integrals such as the one above? The answer to this question is given by a generalization of a procedure known as Leibniz's rule. Stated in the general form in which we need it, this rule says that the derivative of any expression of this form:

$$G(x) = \int_{h(x)}^{k(x)} g(x, y) dy$$

is given by*

$$\frac{dG(x)}{dx} = \int_{h(x)}^{k(x)} \frac{\partial g(x, y)}{\partial x} dy + g[k(x), x] \frac{dk(x)}{dx} - g[h(x), x] \frac{dh(x)}{dx}$$

This one rule will enable us to differentiate all of the integrals which we need for this kind of inventory problem. In our case we have:

$$\begin{aligned} k(x) &= \infty \\ h(x) &= x \\ g(x, y) &= (y - x)f(y) \end{aligned}$$

Therefore, we have

$$\frac{dk(x)}{dx} = 0; \quad \frac{dh(x)}{dx} = 1; \quad \frac{\partial g(x, y)}{\partial x} = -f(y)$$

so,

$$\begin{aligned} \frac{dE.F_z}{dx} &= c + C_u \left[- \int_x^{\infty} f(y) dy + 0 - (x - x)f(x) \right] \\ &= c - C_u[1 - F(x)] \\ &= c - C_u + C_u F(x) \end{aligned}$$

Setting this equal to zero and solving for $F(x)$ gives

$$F(x) = \frac{C_u - c}{C_u}$$

This is the general solution, for any continuous probability distribution of demand. Knowing the probability distribution permits us to determine that value of x for which the given equation holds true. This, then, is the value of x which has the minimum expected total cost. We must note that some conditions must be met to guarantee that this is, in fact, a minimum but they are met here and we will do no more than to state that the second derivative of $E.F_z$ with respect to x must be greater than zero.

Suppose, for example, that the probability distribution of demand for a particular inventory problem of this type were the normal distribution with mean of 100 and standard deviation of 20. If the costs were the same as we used previously, $c = \$100$ and $C_u = \$1000$, we immediately calculate that $F(x) = .9$. Reference to any table of the normal distribution† shows immediately that $F(x) = .9$ is achieved at a point 1.28 standard deviations above the mean. We therefore should order $100 + 1.28(20) = 125.6$ units in order to minimize the expected total cost in this case.

We can equally well formulate this problem in terms of opportunity costs and it may be instructive to do so. Let us put

K_o = opportunity cost of an overstock of one unit

K_u = opportunity cost of an understock of one unit

* See Philip Franklin, *A Treatise on Advanced Calculus* (New York: John Wiley & Sons, Inc., 1940), p. 349.

† See normal tables in Appendix.

The difference between this formulation and the preceding one is that we no longer have a constant cost of cx when we order x units because, since we are thinking in terms of opportunity costs, x units has a cost of zero if the demand level happens to be for x units. Under these circumstances our overstock costs must be calculated exactly as were the understock costs before. Thus, we have:

$$E.F_x = K_o \int_0^x (x - y)f(y) dy + K_u \int_x^{\infty} (y - x)f(y) dy$$

Using our rule for differentiation we find

$$\frac{dE.F_x}{dx} = K_o F(x) + K_u [F(x) - 1]$$

Setting this equal to zero gives

$$F(x) = \frac{K_u}{K_o + K_u}$$

Since this result must be equivalent to the preceding one we can equate the two expressions to find the relationship which must exist between the opportunity costs and the total costs. We find

$$K_o = c$$

and

$$K_u = C_u - c$$

This, of course, is precisely the relationship which does exist between the two costs, as can be seen by looking at the two matrices for the spare part problem of Section 12.

Other costs can usually be incorporated in this kind of formulation in a straightforward way. Suppose, for example, that we have the two costs above, c and C_u , but that in addition we have a salvage value for any unused items of C_s per unit. An obvious extension of the argument leads immediately to the equation:

$$E.F_x = cx - C_s \int_0^x (x - y)f(y) dy + C_u \int_x^{\infty} (y - x)f(y) dy$$

This has the derivative:

$$\frac{dE.F_x}{dx} = c - C_s F(x) - C_u + C_u F(x)$$

Setting the derivative equal to zero gives:

$$F(x) = \frac{C_u - c}{C_u - C_s}$$

The effect of the salvage value is the expected one—to increase the value of x as compared to the same case without any salvage value. Thus, in our normal distribution example suppose each unit has a salvage value of \$50.

We calculate, then, that $F(x) = \$900/\$950 = .947$. Reference to a table of the normal distribution shows that this requires that $x = 100 + 1.62(20) = 132.4$ units.

Carrying costs can introduce a considerable amount of complexity. As a matter of fact, introducing carrying costs as we did in our discrete example resulted in a fair additional amount of computation simply because of the need for discounting future costs and for taking account of the different lengths of time the items which were used were in stock. Generally speaking, this kind of precision is unnecessary for the simple reason that the carrying costs are not known with that order of accuracy. Thus, it is quite common to make a simplification and to assume that only those items which are not used are to be credited with a carrying cost for the time they were in stock. This is a completely reasonable simplification if the total time period involved is relatively short. If, however, a long period is involved it may be necessary to consider other ways of introducing carrying costs. We will not consider this difficulty here. If the simplification is an acceptable one it is straightforward to include the carrying costs. Without repeating the argument it can be seen that the effect of the simplified carrying costs would be to introduce a term of the form $+C_c F(x)$ in the derivative of the expected total costs, where C_c represents the carrying cost. This obviously entails no additional difficulties whatsoever.

15. IMPUTATION OF COSTS

We noted in the first chapter that there is no escaping the existence of the opposing costs in an inventory problem. Ignorance of the costs, or inability to measure them, does not detract in the least from the argument that any decision on inventory level which is actually taken will imply a particular ratio of the opposing costs. As we commented earlier, it is only a question of whether the decision-maker would rather have his assumptions about the costs implicit or explicit. Let us see how this works for the kind of inventory problem we are now considering.

As an example, consider the normal distribution inventory problem used in the preceding section, without any salvage value. Suppose that the decision were made to order 165 units. Since we know the distribution of demand (normal distribution with mean of 100 and standard deviation of 20) we can immediately calculate that this is equivalent to setting $F(x) = .9994$. This follows since 165 units is $65/20 = 3.25$ standard deviations above the mean and reference to a table of the normal distribution shows that this includes .9994 of the total area under the normal curve. We have, then, from our analysis that

$$\frac{C_u - c}{C_u} = .9994$$

or

$$C_u = 1667c$$

In other words, this decision would only be the correct decision if the cost of

further, that the wages of the repairmen are \$20 per day, with time-and-a-half for overtime.

This kind of problem might be approximated by one of our models because of the time required to train a repairman and the difficulty of finding suitable men on short notice. Let us, finally, assume that company policy is to repair every broken phone on the day on which it is reported. If the number of men is insufficient to accomplish this in the regular day they work overtime as much as is necessary. The opposing costs in the problem are clear—a man idle for one day represents \$20 wasted while an understock of repairmen requires the payment of overtime. This is an example of the first kind of model we discussed in Section 14. If x men are hired the expected total cost will be:

$$E.F_x = 20x + 30 \int_x^{\infty} (y - x)f(y) dy$$

The solution is that

$$F(x) = \frac{30 - 20}{30} = .333$$

Looking this up in normal tables shows that $x = 100 - .43(20) = 91.4$ or, rounded off, 91. Thus, the optimal size of the work force is 91 men. It is clear that problems of this kind can easily be solved by inventory models if the necessary assumptions are reasonable ones.

17. GENERAL COMMENTS

The inventory problems we have dealt with in this chapter do not occur as frequently as some of the other kinds of inventory problems we will be considering subsequently. Nonetheless, it is not a rare kind of situation by any means. Further, it is the simplest kind of inventory problem and the logic of the analysis is particularly clear. As such it is a good preliminary to the more difficult kinds of inventory analyses. Generally speaking, any problem of this type can be solved analytically if the necessary information is available concerning costs. There can be difficulties in satisfactorily representing some of the costs but the usual case is relatively straightforward. If the costs cannot be measured the procedures of Section 15 can be used to provide management with some boundaries for the exercise of their intuition and common sense. There are a great number of variants of this kind of inventory problem but almost all of them can be formulated mathematically in one or another of the ways we have discussed in this chapter.

The major defect in the presentation so far has been that we have said nothing about the question of how the information concerning the probability distribution of demand is obtained. We intend to discuss this question to some degree subsequently—without, however, offering any extended treatment, which would require a book in itself. Meanwhile, it should be noted that this is often one of the most difficult questions involved in an inventory

analysis. While it is frequently the case that the required distribution of demand can be obtained it is perhaps more usual that it cannot. Or, at least, it cannot be obtained in the detail required for this kind of analysis. The time and cost of a study to determine the demand distribution can often far exceed the time and cost of a large, and complex, inventory analysis. This being the case, the question immediately arises: What can be done when the demand distribution is not known? When this is the case, of course, we are no longer dealing with an inventory problem under risk, but with one under uncertainty. It is to this subject that we turn in the next chapter.

PROBLEMS

1. (a) A flower store in a small community must stock some orchids to be sold as corsages for the high school prom. Most of the young men will buy gardenias but most of the time a few orchids will be requested. On the basis of past information the store owner estimates the probability distribution of demand to be:

DEMAND	PROBABILITY
0	.05
1	.10
2	.25
3	.35
4	.15
5	.10

Only one order can be placed. The orchids cost the store owner \$5 each and he sells them for \$12.50 each. Delivery of a corsage costs the store owner \$.50 for each corsage. Any orchids unsold after the prom are a total loss. What is the total profit payoff matrix and the optimal order size?

(b) What is the opportunity cost matrix for this problem?

(c) Use opportunity costs and incremental analysis to verify your solution to this inventory problem.

(d) What is the maximum possible worth of additional information to the store owner for this problem?

(e) What salvage value would justify an optimal strategy of ordering one more orchid than found in (a)?

2. (a) A company is making a large boiler installation. A certain automatic monitoring unit is crucial for the operation of the whole system. At the time of the original order spares for this unit can be purchased for \$2000 each. The probability distribution for failures of this unit during the life time of the installation is known to be:

DEMAND	PROBABILITY
0	.35
1	.25
2	.20
3	.15
4	.05

If a spare is needed and is not available the total cost to the company in down time and replacement cost will be \$15,000. Unused spares have no salvage value. What is the total cost matrix and the optimal number of spares to order?

(b) What is the opportunity cost matrix?

(c) What is the maximum possible worth of additional information?

(d) What is the mathematical equation which describes this problem?

3. (a) Demand for an item is normally distributed with mean of 200 units and standard deviation of 40. The overstock cost per unit is \$12 and the understock cost per unit is \$20. What is the optimal order size?

(b) For the item in (a) the decision has been made to order 300 units. Assuming the overstock cost per unit is \$12, what is the imputed understock cost?

static inventory problems under uncertainty

18. GENERAL CHARACTERISTICS

Static inventory problems under uncertainty and under risk have in common the fact that only one order is available. They differ in that risk inventory problems are ones for which we know the probability distribution of demand. Inventory problems under uncertainty are ones for which we do not know this probability distribution. Because of the similarity of the two types of problems the various remarks made in Section 8 concerning costs are equally applicable here and will not be repeated. The frequency and kinds of inventory problems of this type are also similar to the static risk problems.

In the theory and applications of decision theory, the quantitative analysis of decision problems in terms of a payoff matrix, it is customary to distinguish between decisions under risk and under uncertainty. In effect, the distinction has been treated as being virtually a dichotomy. The assumption is made that for decision problems under risk we know the whole probability distribution in question (of demand, for inventory problems). For decision-making under uncertainty it is assumed that we know nothing whatsoever about the probability distribution. Obviously, there are a whole range of intermediate cases which might be considered. These cases, as a matter of fact, are probably of greater practical importance than is either of the two extremes. We don't often have complete knowledge concerning the probability distribution of demand and, on the other hand, we are not often totally ignorant of it. We hope to indicate some methods that may be used in these intermediate cases. We have chosen to include these methods under the discussion of uncertainty—thus, by implication, reserving risk for the case with complete knowledge of the probability distribution. This seems more

in accord with the usual meaning of "uncertainty" than would the other alternative—to put partial knowledge under risk and leave uncertainty for the case of total ignorance.

In this chapter, then, we will try to present first some of the methods which have been suggested for coping with inventory problems under uncertainty in the sense of complete ignorance of the probability distribution. Then we will consider some methods which might be applicable to cases where partial, but not complete, information about this probability distribution is available. Along the way we will pause to cast our vote for the side that maintains that uncertainty in the sense of complete ignorance is very rare. This is one of the sides in a never-ending argument concerning the very meaning of the idea of probabilities. We will not become embroiled in the argument but we will have to mention it in passing.

19. DECISION CRITERIA UNDER UNCERTAINTY

The fact that we are ignorant of the probability distribution of demand in a specific inventory problem does not affect in any way our ability to determine the payoff for each combination of a given order size and a given level of demand. Therefore, we can equally well construct the payoff matrix for inventory problems under uncertainty. The problems special to the case of uncertainty commence when we try to determine which is the best strategy to select, given the payoff matrix. In the case of risk we proceed by calculating the expected payoff for each strategy and then selecting that strategy which has the smallest or largest such expected payoff. In the language of decision theory this part of the analysis is called a decision criterion. It provides the criterion we should use in selecting one from among the available strategies. The whole problem of decision-making under uncertainty, including the special case of inventory decision-making, boils down to one question, what decision criterion should be used? In the case of risk there is a general consensus, supported by strong arguments, that expected values of payoffs should be used. In the case of uncertainty there have been a variety of different criteria suggested, each supported by arguments asserting its plausibility and, sometimes, superiority over the others. We must begin by considering some of the suggested alternatives.

In order to exemplify the various criteria we will utilize a segment of the Christmas tree payoff matrix in total profits as given in Section 9. We do not need the entire matrix so we will use the upper left corner, three-by-three matrix from it:

ORDER	DEMAND		
	1	2	3
1	\$3.50	\$3.50	\$3.50
2	2.00	7.00	7.00
3	.50	5.50	10.50

We assume now that we do not know anything about the probabilities of the various levels of demand. The problem, then, is to select one of the three available strategies. What criterion should we use as a basis for our selection? We will discuss three different criteria which have been proposed. For a far more exhaustive discussion—and a fascinating one—which considers other criteria as well we refer the reader to Chapter 13 of Luce and Raiffa's *Games and Decisions*.*

One major criteria is associated with the name of Abraham Wald. It is called the minimax or maximin criterion. The name comes from the structure of reasoning underlying the criterion. Wald proposes that we should consider each of the available strategies and determine what is the worst that could happen if that strategy were selected. Since the payoff matrix above is expressed in terms of profits the worst that could happen for each strategy is the smallest profit which might be made if that strategy were selected. For our matrix the "worsts" are:

ORDER	WORST OR MINIMUM PAYOFF
1	\$3.50
2	2.00
3	.50

Wald now proposes that we should select that strategy which has the maximum such minimum. In the present case this would result in the selection of the first strategy. In other words, the criterion is to select that strategy which has the maximum among the minima, hence the name, *maximin*. If the original matrix were expressed in terms of costs the worst payoff for each strategy would be the maximum payoff and then we would select that strategy which had the minimum among the maxima, the *minimax* strategy. The reasoning underlying this criterion is that we should always act with the utmost conservatism. Specifically, we always assume that the worst possible thing is going to happen to us. Then we safeguard our position to the maximum extent possible by selecting the strategy which makes this worst as good as possible.

A second criterion has been proposed by Leonard J. Savage.† This is called the regret criterion and is of particular interest to us because it is very closely connected with the opportunity cost matrix. However, let us follow the reasoning which Savage has suggested to afford plausibility to this criterion rather than simply stating the connection of the criterion with the opportunity cost matrix. Savage suggests that the decision-maker really wants to minimize the regret he will experience after he has actually received the specific payoff resulting from his selection of a strategy. In other words, suppose the strategy of ordering three trees is selected. Suppose, further, that the actual demand level which occurs is for one tree. Savage maintains that the

* R. Duncan Luce and Howard Raiffa, *Games and Decisions* (New York: John Wiley & Sons, Inc., 1957).

† Leonard J. Savage, *The Foundations of Statistics* (New York: John Wiley & Sons, Inc., 1954).

decision-maker will experience a regret that he did not select his first strategy and that this regret can be measured by the difference between what he actually made, \$.50, and what he could have made if he had selected the best strategy for the level of demand which actually occurred, \$3.50 in our example. Savage argues that in selecting a strategy the decision-maker should be governed by a desire to minimize this regret which he will ultimately experience. Savage proposes that we begin our analysis of a payoff matrix by constructing the corresponding regret matrix, which shows the regret, so defined, which will result from each combination of a selection of a strategy and the occurrence of a specific level of demand. The method of calculation is precisely the same as that which we used in calculating the opportunity cost matrix for a given payoff matrix and the regret matrix is nothing other than the opportunity cost matrix. In our example we have:

ORDER	REGRET MATRIX		
	DEMAND		
	1	2	3
1	\$0	\$3.50	\$7.00
2	1.50	0	3.50
3	3.00	1.50	0

Savage proposes, finally, to use the minimax criterion on this regret, or opportunity cost, matrix. We have:

ORDER	MAXIMUM REGRET
1	\$7.00
2	3.50
3	3.00

We would select that strategy which had the minimum such maximum regret, the minimax strategy. In our case this is the strategy of ordering three trees.

The final criterion we will consider here is of greater age than the preceding ones and has been the subject of continuing and acrimonious debate for many years. It is usually associated with the name of Thomas Bayes and/or that of P. S. Laplace. Bayes originally suggested the criterion but Laplace was responsible for the first extensive use of it.* This criterion is simple to present. The idea is to convert the uncertainty into risk. The argument is that if we are really totally ignorant of the probabilities of occurrence of the various levels of demand then we should assume that all of them are equally likely. In other words, we should assume that all of them have equal probabilities of occurrence. Once this is granted we can proceed to calculate the expected payoff for each strategy and continue exactly as we do for decision-making under risk. In our case we would assume that the

* See the discussion in Lancelot Hogben, *Statistical Theory* (London: George Allen and Unwin, Ltd., 1957), p. 21.

probability of each level of demand was $\frac{1}{3}$. It is easy, then, to calculate the expected payoff for each strategy:

ORDER	EXPECTED PAYOFF
1	\$3.50
2	5.33
3	5.50

We would select that strategy which had the largest such expected payoff, in this case the strategy of ordering three trees.

In our example we see that the Wald criterion dictated the ordering of one tree while the Savage and Bayes criteria each dictated the ordering of three trees. In other examples it often occurs that each criterion results in the selection of a different strategy. This represents an embarrassment of riches! Faced with such contradictory conclusions on the part of the different criteria, what is to be done? Each criterion has its own rationale and it seems to be only a question of taste as to which criterion is used. This is an unsatisfactory state of affairs and we will devote the next section to some further discussion of these criteria.

20. DECISION CRITERIA FOR UNCERTAINTY AND INVENTORY PROBLEMS

Most of the discussion in the literature dealing with the different decision criteria under uncertainty is framed in terms of the general decision problem. This introduces problems additional to those which characterize the majority of inventory decision problems. Thus, one major difference of the general decision problem is that the payoff matrix may be expressed in terms of some general measure of utility whereas we have excluded this possibility for inventory problems as being rare and atypical. Measurement of utility involves many difficulties and there is a large literature dealing with the suitability of various methods for measuring utility. Among other things, it turns out that we are often not entitled to take differences between utilities as being meaningful. If this is true it would follow immediately, for example, that the Savage regret criterion would not be admissible as a decision criterion for such a payoff matrix. Under these circumstances comparisons among the different decision criteria would be quite different than they would be when we confine our attention to the simpler case where payoffs are measured in dollars. Therefore, we will not attempt to recapitulate any of the conclusions reached in the attempt to compare decision criteria for the general decision problem. Rather, we will consider some of the implications of their use explicitly in terms of inventory decision problems.

When looked at in this way there appears to be a reasonable basis for rejecting some criteria. Thus, consider the Wald criterion. When this criterion is applied to any inventory decision problem of the type of the Christmas tree problem, where the payoffs are expressed in profits, it will generally select that order which is equal to the lowest level of demand. If,

however, we apply the same criterion to the corresponding opportunity cost matrix it will generally select that order which is equal to the highest level of demand. Now, we maintain that this is an inadmissible contradiction in the case of an actual inventory problem payoff matrix. The total cost or profit payoff matrix and the opportunity cost matrix are merely alternative ways of representing the same real problem. We have tried to show previously that the economic logic underlying the two presentations is the same. In other words, the argument which Savage developed to justify the regret matrix in the case of a general decision problem is unnecessary here because the regret matrix is identical with the opportunity cost matrix, which is simply an alternative way of expressing the selfsame problem the original payoff matrix represents. Therefore, we submit that any valid decision criterion should lead to the same selection of strategy whenever it is applied to these two equivalent representations of the same problem. We are suggesting, then, a criterion to be used in the selection of a decision criterion. Any decision criterion is inadmissible which does not select the same strategy when applied to the opportunity cost matrix as it does when applied to the payoff matrix. This suggested criterion is a powerful one because it eliminates not only the Wald criterion but also another well-known decision criterion not discussed here, the Hurwicz criterion. We think that the economic logic of inventory decision problems justifies the application of our criterion and we will therefore not use the Wald criterion.

Both the Savage regret criterion and the Bayes criterion meet this condition. We have already seen that the use of expected values will lead to the same selection of a strategy whether applied to the payoff matrix or to the opportunity cost matrix. The Savage criterion meets our condition because, in effect, it says that the payoff matrix should never be analyzed directly. Instead, it must be converted to the opportunity cost matrix before we consider what strategy to select. This leaves us with two possible decision criteria. In our example of the preceding section these two criteria led to the selection of the same strategy. Whenever this occurs there is no problem. However, whether it occurs or not is a coincidence and we would like to have some way to decide which of these two criteria is more appropriate. Unfortunately, no such argument exists and it is necessary to conclude that it is a question of taste whether the Savage or the Bayes criterion is used. Neither one of them is wholly satisfactory. One major objection against the Savage criterion is that there seems to be no good reason why the decision-maker should concentrate all of his attention on the worst that can happen to him if he selects each strategy. Why not give some weight to the better things that might happen as well? The Bayes criterion gives equal weights to all possible payoffs for each strategy but it is also subject to a severe criticism. This is that if we are truly totally ignorant of the level of demand then we don't know, to take our previous example, that only three different levels of demand are possible. Maybe a demand of four or five could occur. But if we add these possible demand levels to our payoff matrix and again use the Bayes criterion we will not generally end up with the selection of the same strategy. The reader can verify this for himself by adding two columns to the payoff

matrix we used as an example in the last section. Thus, neither of the two available criteria can be considered to be ideal for its purpose.

All of this discussion has been predicated on the assumption of total ignorance of the probabilities of the various possible levels of demand. If we have knowledge about these probabilities we would be dealing with the case of risk, which is much more straightforward with regard to the decision criterion. Now, as a matter of fact, we think that a strong case can be made for the proposition that the case of uncertainty is quite rare and that we can usually convert problems which seem to be under uncertainty into ones under risk. This is the subject of our next section.

21. OBJECTIVE AND SUBJECTIVE PROBABILITIES

One of the reasons that decision-making under uncertainty has received the emphasis in the literature that it has is that the predominant school of statisticians and probability theorists adheres to a definition of probability which automatically makes the case of uncertainty quite a common one in practice. This predominant school is called the objectivist school in probability theory. The objectivists maintain that the only time that the ideas of probability theory can be legitimately applied is when it is possible to determine the relevant probabilities from a frequency count or the equivalent. Thus, we can say that the probability of a flipped coin coming up heads is $\frac{1}{2}$ because we can flip a coin a great many times and count the number of heads. We can talk about the probability that a male picked at random is color-blind because we can make a frequency distribution showing the numbers of men who are, and who are not, color-blind. And we can talk about the probability of a particular level of demand only when we can count past levels of demand or do the equivalent by means of mathematical analysis. Now, since there are a great many kinds of practical cases where we simply cannot count demand levels in the past or accomplish the same thing by mathematical analysis it follows that, according to the objectivists, the corresponding inventory problems must be under uncertainty. This is so by definition, according to them. There simply is no meaning attached to the word "probability" when used in such a context.

Opposed to the objectivist school is the subjectivist school. The subjectivists agree with the objectivists as far as the latter go. In other words, the subjectivists agree that whenever a basis for a frequency count exists it should be used and that probabilities should be determined by this method when it is possible. The disagreement arises in that the subjectivists propose to use probabilities in many cases where there is no possibility for any kind of a frequency count. Specifically, they maintain that probabilities, in addition to being measures obtained from frequency counts, are measures of subjective degrees of belief, hence the name of the school. In this sense, a man who believes that something is certain to happen is ascribing a probability of one to its happening. If he believes that something else is impossible he is ascribing a probability of zero to it. The intermediate probabilities measure

the varying degrees of conviction that the given event will occur. Roughly speaking, a subjective probability of $\frac{1}{2}$ indicates the belief that the event in question is as likely to occur as not, and so forth. Subjectivists have suggested several methods for measuring such degrees of belief and they propose to use them as probabilities when nothing else is available.

The debate between these two schools has a flourishing literature and many arguments pro and con. We do not propose to summarize these arguments here nor do we intend to add to them. One or two comments suggested by typical situations encountered in practice will be sufficient for our purposes. The difficulty in adhering to the objectivist position in practice is that too often there is simply no basis for determining the relevant probabilities in accordance with objectivist canons. The net effect of this is that too many important decision problems cannot be meaningfully analyzed by modern quantitative methods because of the difficulties we have seen with regard to decisions under uncertainty. We feel justified in concluding this because the application of such a criterion as Savage's, for example, to practical inventory problems leads to selections of strategies which are simply not in accord with common sense. Of course, the objectivist's answer to this statement is that our comment is irrelevant; if there is no legitimate basis for resolving a given problem, then it simply can't be resolved. We reject this answer because it ignores a salient feature of most practical business decision problems. This is that there is rarely a case where the relevant company executives are really completely ignorant of the probabilities. To assume that they are is to ignore all of the insights and understanding which can come from experience based on the exercise of good judgement. Typically, the relevant executives know a very great deal about possible demand levels even in cases where there is no basis whatsoever for a frequency count approach. The subjectivists at least offer methods—admittedly imperfect—by means of which we can try to present this executive knowledge in quantitative form so that it can be utilized in the analysis of the executives' decision problems. This appears to us to be sufficient practical reason for taking a stand beside the subjectivists in the debate on the meaning of probabilities.

How, then, might we utilize the subjectivist approach in the attempt to resolve an inventory problem under uncertainty? Let us use the example of Section 19 again. Suppose that there is no basis for estimating the probabilities of the various possible demand levels from past historical data. There are three possible levels of demand and we want to determine the decision-maker's subjective evaluation of the probability of their occurrence. Suppose we present the decision-maker with the following two hypothetical decision problems:

STRATEGY	I DEMAND	
	1	2
A	k	k
B	1	10

STRATEGY	II DEMAND	
	1	3
C	k'	k'
D	1	10

The payoffs given for strategies *B* and *D* are arbitrary except that they should be the same for both hypothetical decisions and they should be of the same order of magnitude as the payoffs in the actual decision problem. This last qualification is to avoid any confounding of the problem of uncertainty with the problem of the utility of money to the decision-maker. Finally, the two payoffs must be selected so that there is a relatively large difference between them.

Now, let us consider the intention and the implications of these hypothetical decision problems. Consider I as an example. Only two of the possible levels of demand can occur in this decision problem. If the decision-maker selects strategy *A* he will receive a payoff of *k* dollars no matter which of them occurs. If he selects strategy *B* he will get the indicated payoff, depending on which of the two demand levels occurs. Let us see what happens as we vary *k*. It is certainly the case that if $k = 0$ the decision-maker would prefer to select strategy *B* since his payoff would always be larger for this strategy. Similarly, if $k = 11$ the decision-maker would always select *A* because his payoff would always be larger. Therefore, as we vary *k* from 0 to 11 there must be some value at which the decision-maker is indifferent between the two strategies. Suppose, for example that he tells us that he is indifferent between *A* and *B* when $k = 8$ and that he is indifferent between *C* and *D* when $k' = 4$. We are now prepared to reconstruct the probabilities of the various demand levels implied by his statements. Since he is indifferent between his strategies for $k = 8$ we reason that for $k = 8$ the expected values of the two strategies must be equal. Therefore,

$$P_1 + 10P_2 = k = 8$$

But, $P_1 + P_2 = 1$, as far as I is concerned, so we calculate that $P_2 = 3.5P_1$. Proceeding similarly for II we find that $P_3 = \frac{1}{2}P_1$. We take these as the values of P_2 and P_3 , expressed in terms of P_1 . Now, returning to the original decision problem, we know that the sum of the probabilities of the three levels of demand must equal one. But their sum in terms of P_1 is $P_1 + 3.5P_1 + .5P_1 = 1$. Solving this we find that $P_1 = .2$ and, hence, that $P_2 = .7$ and $P_3 = .1$. These then are the estimates of subjective probabilities of level of demand which can now be used to calculate the expected values of the various strategies. In short, by this procedure we have converted the original problem under uncertainty into one under risk.

This general procedure is a useful one and it can often be utilized in converting problems under uncertainty into ones under risk. However, it is by no means invariably successful. Several assumptions underlie its applicability and these may break down in specific problems. It is always possible to check the consistency of the estimates of the probabilities by utilizing other hypothetical decision problems involving other combinations of the possible levels of demand. For example, in our case we could check the validity of our calculated values by offering the decision-maker another hypothetical decision problem of the same form based on demand levels two and three. If his new statement of the point at which he would be indifferent is in accord with what would be calculated from the values of the probabilities we may

consider the estimates to be valid. If there is a contradiction it may be necessary to point it out to the decision-maker so that, aware of an inconsistency in his behavior, he can rectify it. For an alternative procedure for the estimation of the subjective probabilities we refer the reader to the book of Leonard J. Savage, previously cited.

22. PARTIAL INFORMATION

So far we have been considering only the kind of decision problem which the objectivist would characterize as being under complete uncertainty. We have tried to show that from another point of view, the subjectivist one, even this kind of problem can be converted to one under risk. Far more frequent, however, is the situation in which some information is available about the distribution of demand—from an objectivist point of view—but not enough to permit us to deduce the probabilities of each possible level of demand. In short, we have partial information. What, if anything, can be done under these circumstances?

Before we undertake to discuss what might be done, the magnitude of the problem involved should be stressed. It is easy enough to write down a mathematical expression which represents the probability distribution of demand and then to construct our analyses on the basis of this distribution. In practice, however, it may be virtually impossible to determine the actual distribution of demand. Most forecasts of demand undertake only to estimate the average level of demand and there are a nondenumerable infinity of probability distributions which have the same average. Yet the complexities and perplexities of the forecasting problem are notorious. In other words, if we cannot succeed in achieving good predictions of average demand, how can we hope to achieve the very much more difficult description of the entire probability distribution? But our analysis of inventory problems under risk requires just this. While this can be accomplished in some kinds of problems it is equally true that it cannot be accomplished in many other situations. It is the latter case that we must consider now.

In order to develop methods for dealing with the case of partial information we must know what we mean by partial information. We could have partial information about the level of demand in an enormous number of different ways. We might know, as examples, that demand would be confined to some given range, that demand would be less than some given maximum amount, that the probabilities of lower levels of demand are greater than those for higher levels of demand, that the probability distribution of demand is unimodal, that the probability distribution is symmetrical, that it is skewed to the left, or we might know the average level of demand only, and so forth. A perfect system for the utilization of partial information would include two major components. First, it would offer a graduated series of classes of partial information so that, whatever we knew about demand, we could fit our knowledge of demand in a specific problem neatly into one of the classes. Second, it would then show exactly how the given

class of information could be optimally utilized in resolving the specific inventory decision problem. Since extremely little progress has been made in this area it is obvious that there is no question of presenting such a system. Nonetheless, these desiderata suggest the framework for the discussion that follows.

As a generalization we can say that the vaguer and more qualitative our partial information, the less possibility exists, granted our present state of knowledge, of utilizing the information in a meaningful way. Mostly, then, we will have to confine our attention to cases where the partial information can be expressed quantitatively. In other words, while we cannot characterize completely the probability distribution of demand we do know some quantitative things about it. A natural way in which partial information may be expressed is in terms of the moments of the demand distribution—we may know some of the moments without knowing all of them. The r th moment of a probability distribution is the expected value of the r th power of the variable:

$$\text{Moment of order } r = \mu_r = \int_0^{\infty} y^r f(y) dy$$

More usually, however, we are interested in the moments around some point, x . These are defined by:

$$\text{Moment of order } r \text{ around } x = m'_r = \int_0^{\infty} (y - x)^r f(y) dy$$

The particular point, x , which we are most often interested in using is the mean of the distribution, $\bar{z}$. Moments around the mean are, then,

$$\text{Moment of order } r \text{ around the mean} = m_r = \int_0^{\infty} (y - \bar{z})^r f(y) dy$$

The mean, of course, is simply the first moment, μ_1 . We can derive relationships between the moments about the mean and the corresponding moments, μ_r :

$$m_1 = 0$$

$$m_2 = \mu_2 - \mu_1^2 = \sigma^2$$

$$m_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3$$

$$m_4 = \mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4$$

Finally, there is another kind of moment, the absolute moments about a point x , which is less commonly used but which is particularly useful for our purposes. The absolute moments around a point x are defined by

$$\text{Absolute moment of order } r \text{ around } x = \lambda'_r = \int_0^{\infty} |y - x|^r f(y) dy$$

and

$$\text{Absolute moment of order } r \text{ around the mean} = \lambda_r = \int_0^{\infty} |y - \bar{z}|^r f(y) dy$$

It can be seen immediately that for all even-order moments we have $m_{2r} = \lambda_{2r}$. This follows since any even power of a difference is always positive, exactly as is the absolute value. For the odd-order moments, of course, the equality does not hold—the absolute moment of order $(2r + 1)$ is always greater than the corresponding m_{2r+1} .

Most persons who have taken a course in statistics will be familiar with the first moment, the arithmetic mean, and with the second moment about the mean, the variance. They may be less familiar with the third and fourth moments about the mean because these are not so often calculated in practice since it is difficult to obtain satisfactory estimates of them unless there is a sizable amount of data available. Nonetheless, the third and fourth moments about the mean are of fairly common use. The third moment about the mean is useful in characterizing the divergence from symmetry, or skewness, of a distribution. It is convenient for this purpose to use

$$\gamma_1 = \frac{m_3}{\sigma^3}$$

as a measure of skewness. This quantity is zero for a symmetrical distribution and is negative if the longer tail is left of the mean and positive if the longer tail is to the right of the mean. The fourth moment is useful in characterizing the peakedness, or kurtosis, of a distribution. For this purpose we often use

$$\gamma_2 = \frac{m_4}{\sigma^4} - 3$$

as a measure of kurtosis. The reason for the subtraction of 3 is that when γ_2 is so defined the normal distribution has $\gamma_2 = 0$.

These first four moments supply the basis for the well-known system developed by Karl Pearson to represent practical probability distributions.* Pearson used a differential equation which defines a whole class of possible probability distributions. The specific distribution of this class which would best represent a particular given distribution was completely determined by knowing the first four moments around the mean of the given distribution. In the overwhelming majority of cases the Pearsonian system provides a satisfactory representation of any practical distribution. This fact indicates that the first four moments about the mean of a distribution summarize the greater part of the information about the distribution. Most probability distributions have an infinity of moments but there is a principle of diminishing returns operating with regard to the additional information about the distribution which is supplied by each successive moment.

These remarks serve to indicate that one method of characterizing partial information is in terms of the moments about the mean of the demand distribution. The more moments we know, the more we know about the distribution. It turns out that there is virtually nothing we can do if we do not know at least the mean and the variance of the distribution—the first two

* For a discussion see Maurice G. Kendall, *The Advanced Theory of Statistics*, vol. 1 (London: Charles Griffin & Co., Ltd., 1947).

moments. If we do know these, however, there is a considerable amount that we can do with regard to the resolution of inventory problems. And, of course, we expect that knowledge of further moments should permit us to do better. The next three sections are devoted to the development of methods for utilizing partial information in this sense.

It is unfortunately the case that not all of the information we might have about a probability distribution can be simply summarized in terms of the moments of the distribution. For example, we might know that the demand distribution was symmetrical about the mean—the probability of a demand of $(\bar{z} + k)$ equals the probability of a demand for $(\bar{z} - k)$, for all k . We have already mentioned that our measure of skewness, γ_1 , will be zero for a symmetrical distribution. However, γ_1 can equal zero without the distribution being symmetrical. Now, it may well be the case that in practice if we think that the demand distribution is symmetrical we really mean only that $\gamma_1 = 0$. However, it is worth noting that to know that a distribution is really symmetrical requires that we know that all the odd-order moments about the mean are equal to zero. In other words, symmetry is not a characteristic of a distribution which can be expressed in terms of one or two moments—it requires an infinity of them. It is even worse with regard to unimodality of a distribution. Unimodality means that a demand distribution, for example, only has one level of demand which has the maximum probability. The probabilities of all levels of demand to the right and to the left of this modal level decrease steadily. It is a particularly reasonable assumption that a demand distribution should be unimodal since most frequently occurring distributions are unimodal—the normal, binomial, and Poisson distributions are all examples. Yet there is no way in which we can characterize the fact of unimodality by means of the moments of the distribution. Our measure of skewness, γ_1 , will indicate how far the mode is from the mean of the distribution but the fact of unimodality escapes definition in terms of the moments. We will endeavor, in the following sections, to indicate how some of these kinds of partial information can be utilized.

Before we proceed to a discussion of the procedures for utilizing partial information we should title this approach in terms of its parallels with problems in statistics. Much use is made in applied statistics of methods which are called non-parametric. The moments of a probability distribution are also called its parameters so non-parametric methods are ones which do not depend on the parameters of the relevant distributions. By extension the same title is sometimes given to methods which, while depending on knowledge of some of the underlying parameters, do not depend on knowing all of them. With the same extension we could call our approach to inventory problems non-parametric. However, in strict logic this is an incorrect description since knowledge of some parameters is required. We will therefore call these methods distribution-free, since they do not depend on a knowledge of the underlying probability distribution, as do the methods developed in the preceding chapter.*

* See Sidney Siegel, *Non-Parametric Statistics for the Behavioral Sciences* (New York: McGraw-Hill Book Co. Inc., 1956), p. 3.

23. DISTRIBUTION-FREE INVENTORY ANALYSIS: THE TCHEBYCHEFF INEQUALITY

The methods developed in the preceding chapter utilize the probability distribution of demand in order to deduce the optimal amount of stock to order. When we confine ourselves to a knowledge of only the average demand and the variance of demand we are obviously giving up a very great amount of information, as compared to the complete knowledge of the demand distribution. This is implicit in the fact that there are a nondenumerable infinity of probability distributions with the same average and variance. If all we know is the average and the variance then any one of this infinity of possible distributions may be the actual one. We cannot, then, determine the optimal inventory policy in terms of any specific distribution. Instead, we must find that policy which will be best in terms of the whole infinite class of possible distributions. In order to be able to do this we must have a starting point in the sense of some kind or kinds of probability statements which we can make which will be true for every specific probability distribution in the class with the same mean and variance. We will designate the mean and variance by $\bar{z}$ and s^2 , respectively, in accordance with the usual notation for these parameters. The standard deviation of the class of probability distributions, equal to the square root of the variance, will then be simply s .

The starting point we will use is a well-known inequality of probability theory, known as Tchebycheff's inequality. This inequality states that for any probability distribution whatsoever there is a simple relationship expressing the probability that the given variable will differ from its mean by some multiple of its standard deviation. Expressed specifically in terms of demand distributions the inequality states that

$$P(|y - \bar{z}| \geq ks) \leq \frac{1}{k^2} \quad \text{where } k > 0$$

where $|y - \bar{z}|$ designates the absolute value of the difference. In other words, positive and negative values of $(y - \bar{z})$ are both included in the statement. Translated, this inequality tells us that the probability that demand will differ from the mean by more than ks is always less than or equal to $1/k^2$. This is easily proved. By definition of the variance we have that

$$s^2 = \sum_{i=0}^{\infty} (i - \bar{z})^2 P_i$$

Since all the terms in the summation are positive we can only decrease the sum if we delete some of the terms. Let us delete all terms for which $(i - \bar{z})^2 \leq t^2$, where t is an arbitrary positive number. Let us denote the remaining terms by a summation with an asterisk. Clearly,

$$s^2 \geq \sum_{i=0}^{\infty *} (i - \bar{z})^2 P_i$$

* We follow William Feller, *An Introduction to Probability Theory and its Applications*, 2nd ed. (New York: John Wiley & Sons, Inc., 1957), p. 219.

Each of the remaining terms is, by definition, $> t^2$ so

$$s^2 \geq t^2 \sum_{i=0}^{\infty} P_i$$

But $\sum_{i=0}^{\infty} P_i$ is precisely the probability that $(i - \bar{z})^2 \geq t^2$, since we have deleted all other terms. Therefore

$$t^2 P(|i - \bar{z}| \geq t) \leq s^2$$

or
$$P(|i - \bar{z}| \geq t) \leq \frac{s^2}{t^2}$$

Substituting $t = ks$ we get the stated relationship

$$P(|y - \bar{z}| \geq ks) \leq \frac{1}{k^2}$$

Now, how can this relationship be used in an inventory problem formulation? We intend to develop some better estimates of the probability in question but it will be worth while to illustrate how we can use this basic inequality directly. Let us take an example similar to some in the preceding chapter but with a simpler set of costs than any of those. Suppose that the cost of the item is c for each unit and that the cost of understock is a constant, K , independent of how many units are short. If we formulate this problem as in the last chapter we have:

$$E.F_z = cx + K \int_x^{\infty} f(y) dy$$

This is minimized by $f(x) = c/K$. Suppose, for example, that the distribution of demand is normal with mean of 50, standard deviation of 5, and that the costs are $c = \$5$ and $K = \$1000$. It follows that $f(x) = .005$. In order to use the standard tables for the normal distribution it must be remembered that the ordinate, $f(x)$, undergoes a change of scale when we transform to the standard normal distribution. Thus, to find where $f(x) = .005$ we must find where the standard normal distribution has an ordinate of $.005s = .025$ here. This occurs at a distance of $2.35s$ from the mean. Thus, the optimal amount of stock to order is $50 + 2.35(5) = 61.75$ with a total expected cost of nearly \$318.

Before we try to formulate the same problem by means of the Tchebycheff inequality we should dispose of one question concerning the just-completed analysis. Since the normal distribution has the same ordinate at equal distances from the mean, how do we know that the minimum expected total cost is not minimized at $50 - 2.35(5) = 38.25$ instead of at 61.75? The answer is that for any unimodal distribution, such as the normal distribution, and for the specific problem in question, the minimum expected cost cannot occur at any value less than the mode unless it is the smallest possible level of demand, which would be zero in most cases. This is so because the second

derivative of the expression above is always less than zero for any value less than the mode. In the normal distribution, of course, the mode equals the mean so the specific problem we are considering will never have the optimal level of stock less than the mean as long as a normal distribution is assumed. The only proviso to this statement is that the optimal amount to order may be zero. This would be indicated from our solution whenever c/K is so large that there is no ordinate of the normal curve with that value.

Now let us see how the same problem might be formulated more generally by means of the Tchebycheff inequality. For any probability distribution whatsoever which has a mean of $\bar{z}$ and a standard deviation of s we know that

$$P(|y - \bar{z}| \geq ks) \leq \frac{1}{k^2}$$

The worst possible case, then, would be when the right-hand equality held and this is what we will assume. For any probability distribution of demand the general expression for the expected total cost would be, analogically to the preceding formulation:

$$E.F_x = cx + KP(y > x)$$

For a continuous distribution, our present case, this is equally well

$$E.F_x = cx + KP(y \geq x)$$

Suppose, now, that we wrote our x in the form $x = \bar{z} + ks$. The equation becomes

$$E.F_x = c(\bar{z} + ks) + KP(y \geq \bar{z} + ks)$$

but

$$P(y \geq \bar{z} + ks) = P(y - \bar{z} \geq ks) \leq P(|y - \bar{z}| \geq ks) \leq \frac{1}{k^2}$$

Therefore, we can write

$$E.F_x = c\bar{z} + cks + \frac{K}{k^2}$$

where we assume the worst case, equality, as noted above.* This expression is to be minimized with respect to k . We find

$$\frac{dE.F_x}{dk} = cs - \frac{2K}{k^3} = 0$$

or

$$k = \sqrt[3]{\frac{2K}{cs}}$$

This value of k will minimize the expected total cost for the worst possible case of a probability distribution with mean of $\bar{z}$ and standard deviation of s .

* This is the worst possible case because the expected cost would obviously be smaller if the demand probability distribution were such that the indicated probability was actually less than $1/k^2$.

Assume, now, that we know only that the demand distribution has a mean of 50 and a standard deviation of 5, with the same costs as before. We calculate

$$k = \sqrt[3]{\frac{2(1000)}{5(5)}} = 4.31$$

Therefore, $x = 50 + 4.31(5) = 71.55$ is the optimal order size over any possible probability distribution of demand with the given mean and standard deviation. This is, in effect, a minimax solution—but it is over probability distributions, not over strategies.

Thus, if the given demand distribution were actually normal and we didn't know it our ignorance would result in the order of 71.55 units instead of the optimal 61.75. We would think the expected total cost was

$$5(71.55) + \frac{1000}{(4.31)^2} = \$411.60$$

whereas it would actually be only \$357.75, since a stock of 71.55 with the given normal distribution would produce a probability of nearly zero of going out of stock. Thus the actual extra cost of our ignorance would only be about \$40 or about 12.6% of the optimal \$318. Of course, this calculation assumes the demand distribution is actually normal. Our distribution-free analysis is predicated on the assumption that we do not know nearly so much about demand and, under this assumption, 71.55 units is the optimal order size.

We have reasoned directly from the Tchebycheff inequality, using the fact that

$$P[(y - \bar{z}) \geq ks] \leq P[|y - \bar{z}| \geq ks] \leq \frac{1}{k^2}$$

However, when we are interested only in the probability of deviations from the mean in one direction, as here, a better upper limit of the probability is available. By other means we can derive*

$$P[(y - \bar{z}) \geq ks] \leq \frac{1}{k^2 + 1}$$

Using this inequality, and assuming equality as the worst case, we would have

$$E.F_z = c\bar{z} + cks + \frac{K}{k^2 + 1}$$

We find

$$\frac{dE.F_z}{dk} = cs - \frac{2kK}{(k^2 + 1)^2} = 0$$

From which we obtain the equation

$$k^4 + 2k^2 - \frac{2K}{cs}k + 1 = 0$$

* See Harald Cramér, *Mathematical Methods of Statistics* (Princeton: Princeton University Press, 1946), p. 256, Prob. 5.

This equation has the root $k = 4.15$. Therefore, $x = 50 + (4.15)(5) = 70.75$ is the optimal order size. We would ordinarily use this solution rather than the preceding one but the difference between the two is relatively small here and the easier calculations of the earlier method may often be satisfactory approximations to these more exact calculations.

24. DISTRIBUTION-FREE ANALYSIS: SOME OTHER INEQUALITIES

The analysis of the preceding section is the best we can do if we know nothing whatsoever about the probability distribution of demand other than the mean and the standard deviation. However, if we have more information we can improve our selection of the optimal policy. Suppose, for example, that we know that the probability distribution is symmetrical. In this case the mean of the distribution equals the median and the existence of symmetry can be characterized by $f(\bar{z} + k) = f(\bar{z} - k)$, for all k . Under this condition it is evidently true that

$$P[(y - \bar{z}) \geq ks] = P[(\bar{z} - y) \geq ks] = \frac{1}{2} P(|y - \bar{z}| \geq ks) \leq \frac{1}{2k^2}$$

If we repeat our previous analysis, using this upper limit for the probability, we find

$$k = \sqrt[3]{\frac{K}{cs}}$$

In our example this gives $k = 3.42$. This leads to an order of $x = 50 + 5(3.42) = 67.1$ units with an expected total cost of \$378.50. If the normal distribution actually obtained, the expected cost would be, for this order, \$335.80 or about 6% more than optimal. This represents the cost of our ignorance. It can be seen that the additional information that the distribution is symmetrical has decreased our order size.

The Tchebycheff inequality was proved above for discrete distributions. It is much easier to prove—and in a very much more general form—if we use continuous distributions. We can make the most direct statement in terms of the absolute moments of the distribution. We know that

$$\lambda_r = \int_0^{\infty} |y - \bar{z}|^r f(y) dy$$

Let us take out all the y 's such that $|y - \bar{z}| < t$ and designate the set of remaining y 's by S . Since all of the terms removed are positive it follows that the integral can only be decreased if we integrate only over S . Thus,

$$\lambda_r \geq \int_S |y - \bar{z}|^r f(y) dy$$

But every $|y - \bar{z}|$ in this last integral is $\geq t$ so

$$\int_S |y - \bar{z}|^r f(y) dy \geq t^r \int_S f(y) dy$$

The integral $\int_S f(y) dy$ is integrated over precisely those y 's for which $|y - \bar{z}| \geq t$. In other words,

$$\int_S f(y) dy = P(|y - \bar{z}| \geq t)$$

Therefore,

$$\lambda_r \geq t^r P(|y - \bar{z}| \geq t)$$

or

$$P(|y - \bar{z}| \geq t) \leq \frac{\lambda_r}{t^r}$$

Now, if we put $t = k\sqrt[r]{\lambda_r}$ we get

$$P(|y - \bar{z}| \geq k\sqrt[r]{\lambda_r}) \leq \frac{1}{k^r}$$

This is the desired generalization of the Tchebycheff inequality. Substituting $\lambda_2 = m_2 = s^2$, we see that we get exactly the original Tchebycheff inequality. But our present statement permits us to utilize knowledge of any absolute moment which we may know. If we set up our expected payoff equation using this upper limit to the probability as the worst possible case we find:

$$E.F_x = c\bar{z} + ck\sqrt[r]{\lambda_r} + \frac{K}{k^r}$$

which is minimized by

$$k = \sqrt[r+1]{\frac{rK}{c\sqrt[r]{\lambda_r}}}$$

The K in these last two equations is, of course, the fixed understock cost. Suppose, in our previous example, that we know that $\lambda_4 = m_4 = 1875$. This would be the value if the demand distribution in question were actually normal. Putting in the values in our expression for the minimizing k gives

$$k = \sqrt[5]{\frac{4000}{5(6.58)}} = 2.61$$

The optimal order size would therefore be $x = 50 + 2.61(6.58) = 67.17$. If the distribution were actually normal the expected cost would be, with this order size, about \$336. The cost of our ignorance would therefore be about 6% of the optimal cost.

Here, again, we might know that the distribution is symmetrical. Exactly as before, this knowledge would cut the upper limit of the probability in half. This would give for the minimizing k :

$$k = \sqrt[r+1]{\frac{rk}{2c\sqrt[r]{\lambda_r}}}$$

Assuming this in our example gives $k = 2.27$ or $x = 50 + 2.27(6.58) = 64.94$.

This would give a total expected cost of about \$326 or a cost of ignorance of about 2.5%.

Assuming that we know both $\lambda_2 = m_2$ and $\lambda_4 = m_4$ we could use either of the two inequalities we have derived. Which is better? Obviously, that one is better which gives the smaller total expected cost. Generally it will be the one using the higher moment around the mean but this will not always be the case. As an instance, in our example if everything else remained the same but we change the fixed cost of an understock to any amount greater than \$58,100 we will find that the original Tchebycheff inequality is better for our purposes. A mathematical expression can be derived to show the conditions under which one is better than the other but it has so many fractional exponents that it is easier simply to check both possibilities if a comparison is required.

The question can be raised: Could we not somehow utilize both moments, when we know them, and achieve a better estimate of the upper limit of the probability than we get from either one of them singly? This question was answered affirmatively by Cantelli in 1910.* We will not derive the expression here but we will state it. Knowing the second and fourth absolute moments (equal to the second and fourth moments about the mean), we have:

$$P(|y - \bar{z}| \geq ks) \leq \frac{\lambda_4 - \lambda_2^2}{\lambda_4 + k^4\lambda_2^2 - 2k^2\lambda_2^2} = \frac{\lambda_4 - s^4}{\lambda_4 + k^4s^4 - 2k^2s^4}$$

This can also be written in terms of our measure of kurtosis, γ_2 , as

$$P(|y - \bar{z}| \geq ks) \leq \frac{\gamma_2 + 2}{(k^2 - 1)^2 + \gamma_2 + 2}$$

If we use this upper limit for the probability in our equation for the expected cost we can take the derivative and set it equal to zero as before, thus finding the optimal value of k . The advantage of writing the expression for the probability in terms of γ_2 is that we can often summarize our knowledge of the distribution with regard to the fourth moment directly in terms of γ_2 rather than in terms of λ_4 or m_4 . For the normal distribution we know that $\gamma_2 = 0$. When we assume a demand distribution is normal we are assuming a knowledge of the values of all the even-order moments about the mean plus the fact that all the odd-order moments about the mean are zero. To illustrate the use of this inequality let us assume that we know the costs, the mean, and standard deviation as in our example plus the fact that for the demand distribution $\gamma_2 = 0$. This is still a great deal less than assuming that the distribution is normal. Proceeding as before, we find that the minimizing value of k is given as the root of this equation:

$$\frac{k(k^2 - 1)}{[(k^2 - 1)^2 + 2]^2} = \frac{cs}{8K}$$

* See Maurice Fréchet, "Recherches Theoriques Modernes sur le Calcul des Probabilites," *Traite du Calcul des Probabilites et des ses Applications*, Tome I, Fascicule 3, ed. by Emile Borel (Paris: Gauthier-Villars, 1950), p. 144.

This expression is too complicated to solve explicitly for k but it is not difficult to obtain a good approximation quickly by constructing a table showing how the left-hand side of the equation varies with k :

k	$\frac{k(k^2 - 1)}{[(k^2 - 1)^2 + 2]^2}$
2	.04959
2.5	.01503
3	.00551
3.5	.00238
4	.00116

For our example, $cs/8K = .003125$. We see from our table that k must fall between 3 and 3.5. Interpolating and trying a couple of intermediate values we can quickly determine that $k = 3.33$. This leads to $x = 50 + 5(3.33) = 66.65$. If the normal distribution actually obtained this order would produce an expected total cost of about \$334, or a cost of ignorance of about 5% of the optimal cost. Here, again, we might know in addition that the demand distribution was symmetrical. If so, we can improve our upper limit for the probability by dividing by two, as we have done before. We leave it to the reader to verify that in this event and for our example, $k = 2.93$ and, therefore, $x = 64.65$. As compared to an actual normal distribution, the cost of our ignorance here would be only about 2.2% of the optimal cost.

Let us now turn our attention to the question of what can be done when the frequently plausible assumption of unimodality can be made. As we noted earlier, we can not characterize unimodality in terms of the moments of the distribution. However, if we have reason to assume that the distribution is unimodal then the third moment about the mean, m_3 , is a measure of where the mode is. The measure of skewness, γ_1 , is simply this third moment made dimensionless by dividing by the third power of the standard deviation. In order to get an intuitive feeling for this measure of skewness it is worthwhile to follow another route. In a symmetrical unimodal distribution, like the normal distribution, the mean equals the mode equals the median. It seems natural, then, to measure skewness by the distance of the mode or the median from the mean. Pearson suggested such a measure of skewness, w , where

$$w = \frac{\bar{z} - \text{mode}}{s}$$

The advantage of this Pearsonian measure of skewness is that it is very simple and has an evident meaning in terms of the probability distribution. Now, for the Pearsonian system of probability distributions it can be shown that

$$w = \frac{\gamma_1(\gamma_2 + 6)}{2(5\gamma_2 - 6\gamma_1^2 + 6)}$$

When γ_1 and γ_2 are small this is approximately $w = \frac{1}{2}\gamma_1$.* This shows the

* See Harald Cramér, *Mathematical Methods of Statistics* (Princeton N.J.: Princeton University Press, 1946), p. 184.

relation between Pearson's measure of skewness and the third moment about the mean, m_3 . Thus, one can either estimate the measure of skewness by calculating the third moment about the mean or one can estimate it from a knowledge of the mean, the standard deviation, and the mode—plus, of course, the assumption that the distribution is unimodal. Knowing w , it can be shown that*

$$P(|y - \bar{z}| \geq ks) \leq \frac{4}{9} \frac{(1 + w^2)}{(k - |w|)^2} \quad \text{for } k > |w|$$

For a normal distribution $w = 0$. Let us assume that in our example we know the mean, the standard deviation, that the distribution is unimodal, and that $w = 0$. Under the assumption that $w = 0$, the minimizing value of k is given by

$$k = \sqrt[3]{\frac{8K}{9cs}}$$

In our example this gives a value of $k = 3.29$ or $x = 66.45$. Once again we are entitled to divide our upper limit by two if we know that the distribution is symmetrical. This would give $k = 2.61$ or $x = 63.05$, almost as small an order as is given by assuming normality.

This inequality will not produce a lower amount of stock to order unless w is reasonably small. The reason for this can be seen from the minimizing expression for k which results from using this upper limit for the probability:

$$k = \sqrt[3]{\frac{8(1 + w^2)K}{9cs}} + |w|$$

This estimate of k will be larger than that of the Tchebycheff inequality unless w is appreciably less than one in absolute value. Even when w is small the improvement from using this inequality may not be very large. For example, suppose that in our example we know that the mode is 47.5. Then $w = (50 - 47.5)/5 = .5$. Our estimate of k , from the last equation, is now 4.04 with a resulting $x = 70.2$. The use of the Tchebycheff inequality on this example, without knowing the mode, led to $k = 4.31$. Thus, there has been an improvement but not a huge one.

The reason for this modest improvement (except when $w = 0$) is that we are not really using our information correctly. The point is that if we know the mode of the distribution we should use the mode as our central point rather than the mean. Let us designate the mode by M_o and the second moment about the mode by t . We can easily calculate t by means of the relationship: $t^2 = s^2 + (M_o - \bar{z})^2$. Gauss had already shown in 1821 that for a continuous unimodal distribution the following inequality obtained:†

$$P(|y - M_o| \geq kt) \leq \frac{4}{9k^2}$$

* Cramér, *Mathematical Methods of Statistics*, p. 183.

† See Cramér, work and page cited.

We now must express the order size in terms of the mode instead of the mean. We have:

$$E.F_x = cM_o + ckt + \frac{4K}{9k^2}$$

$$\frac{dE.F_x}{dk} = ct - \frac{8K}{9k^3} = 0$$

and

$$k = \sqrt[3]{\frac{8K}{9ct}}$$

Using this inequality on our example with a mode equal to 47.5 leads to $k = 3.17$ and $x = 47.5 + 3.17(5.59) = 65.22$. Employing our information in this way, it can be seen, results in a considerable improvement in the optimal size of the order.

Finally, we may note that our last inequality can be generalized to cases where moments of higher order than two are known. Again, for maximum efficiency we must use the mode as our central point and express the higher moments around the mode rather than around the mean. Thus, for example, suppose we know m_3 and m_4 . Then we can calculate the fourth moment around the mode, t_4 , from the equation:

$$t_4 = m_4 + 4(\bar{z} - M_o)m_3 + 6(\bar{z} - M_o)^2m_2 + -3(\bar{z} - M_o)^4$$

The inequality, then, is*

$$P(|y - M_o| \geq k\sqrt[4]{t_4}) \leq \frac{16}{25k^4}$$

Using this upper limit for the probability in our equation for expected cost we find the minimizing value of k is given by

$$k = \sqrt[5]{\frac{64K}{25c\sqrt[4]{t_4}}}$$

And, of course,

$$x = M_o + k\sqrt[4]{t_4}$$

We will illustrate the use of this inequality on our example. Suppose we have $m_2 = 25$, $m_3 = 125$, and $m_4 = 2000$ with $M_o = 47.5$ and the mean = 50. Putting these values in the expression for t_4 we find $t_4 = 4227$. The minimizing value of k is, then, 2.29 and the optimal order size is 66.00.

These various inequalities provide means for utilizing partial information of the kinds we have discussed.† However, we have only considered one kind of cost equation which assumes a fixed cost for understocks of any

* See Maurice Fréchet, opus cited, p. 162.

† For a thorough treatment of all of these inequalities, and others, we refer the interested reader to the previously cited work of Fréchet. In reference to some of these inequalities he says, on page 156: "Sans grande utilité dans la théorie des probabilités, les formules obtenues peuvent, au contraire, rendre de grands services en statistique du les lois de probabilité ne sont souvent qu'imparfaitement connues."

amount. We must consider in the next section procedures which may be followed if there is a per unit cost of understocks.

25. DISTRIBUTION-FREE ANALYSIS: PER UNIT COST OF UNDERSTOCK

We will suppose, now, that we have an understock cost, C_u , for each unit we are short. We have already seen that this kind of situation results in an expected cost equation of the form

$$E.F_z = cx + C_u \int_x^{\infty} (y - x)f(y) dy$$

which can be minimized over x directly by

$$F(x) = \frac{C_u - c}{C_u}$$

We want to find a distribution-free method of analysis for this kind of problem.

Specific methods are available for the resolution of this problem. However, the presentation of these methods is a rather lengthy matter and we will not undertake it here.* Instead an approximate method is developed in the Appendix. Assuming that $x = \bar{z} + ks$, and using the Tchebycheff inequality, we find that

$$E.F_z \leq c(\bar{z} + ks) + C_u s \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

to a sufficient approximation for our purposes. We want to minimize $E.F_z$ with respect to k . Differentiating and setting the resulting expression equal to zero gives

$$k^4 - \frac{C_u}{c} k^2 - \frac{C_u}{c} k - \frac{C_u}{2c} = 0$$

The positive root of this equation, for specific values of c and C_u , is the desired value of k and, of course, of $x = \bar{z} + ks$. Taking $\bar{z} = 50$, $s = 5$, $c = \$5$, and $C_u = \$50$ we find $k = 3.62$ or $x = 68.1$. For this size order our equation shows that $E.F_z \leq \$420$. For a normal distribution with the same parameters we would find an optimal order size of $x = 56.4$ with an expected total cost of \$294.† If this normal distribution actually obtained our order of 68.1 would have an actual expected cost of only \$341 so, in this case, the cost of our ignorance would be about 16%.

* These methods are developed in the dissertation of David W. Miller, "The Utilization of Partial Information in Decision Problems" (Graduate School of Business, Columbia University).

† In order to calculate this expected total cost and all subsequent similar ones based on the assumption of normality it is necessary to determine partial means of the normal distribution. Let $f(y)dy$ be the normal distribution with mean of $\bar{z}$ and standard deviation of s . Then,

$$\int_x^{\infty} yf(y)dy = \bar{z}[1 - F(u)] + sf(u)$$

where u is the standardized normal deviate corresponding to x , $\left(u = \frac{x - \bar{z}}{s}\right)$.

If we assumed a symmetrical distribution our argument would lead to the equation for k :

$$\frac{4c}{C_u} k^4 - 2k^2 - 2k - 1 = 0$$

For our example this would give $k = 2.69$ or $x = 63.45$.

Our derivation depends on the assumption we made that the amount ordered would be greater than the mean of the demand distribution. Unfortunately, our equation offers no indication of whether this assumption is justified or not. Any feasible values of c and C_u will produce a value of k and, hence, of $x = \bar{z} + ks$, even if the true optimal size of order is less than the mean. As can be seen by consideration of the general form of the original solution of this kind of inventory problem,

$$F(x) = \frac{\bar{C}_u - c}{C_u}$$

the optimal order should be zero for $c = C_u$ and should increase as c decreases. The equation we have derived does not reflect this fact. It will always give an optimal order greater than $\bar{z}$, no matter what the values of c and C_u .

When we assume an order size less than the mean of the demand distribution we find, as indicated in the Appendix, that

$$E.F_x \leq c(\bar{z} - ks) + 2ksC_u + C_us \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

This equation is based on the Tchebycheff inequality. The minimizing value of k is given as the positive root of

$$2 \left(\frac{2C_u - c}{C_u} \right) k^4 - 2k^2 - 2k - 1 = 0$$

Suppose $\bar{z} = 50$, $s = 5$, $c = \$40$, and $C_u = \$50$. We find $k = 1.31$ or $x = 43.45$. If the demand distribution were normal the order size would be about $x = 45.8$.

Let us assume the demand distribution is unimodal and symmetrical. Then we have the equation:

$$E.F_x \leq c(\bar{z} - ks) + 2 C_u ks + \frac{2}{9} C_us \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

This gives the following equation for k :

$$9 \left(\frac{C_u - c}{C_u} \right) k^4 - 2k^2 - 2k - 1 = 0$$

For the data of our preceding example we find $k = 1.46$ or $x = 42.7$. This seems to have moved in the wrong direction but it is an illusion. The first

equation based on $x = \bar{z} - ks$ gives, for our example, $E.F_z \leq \$2775$. The one we have just calculated gives $E.F_z \leq \$2127$. If demand were actually normal, $E.F_z = \$1820$. Thus, the changes in $E.F_z$ are in the right direction.

The conditions which serve to establish whether to use $x = \bar{z} + ks$ or $x = \bar{z} - ks$ are remarkably complicated and will not be given here. In practice, in cases where there is any doubt it is easier to try both and see which has the actual minimum $E.F_z$. When c is small relative to C_u it can be expected that x will be greater than the mean. As the ratio of c to C_u approaches unity the amount to be ordered will approach zero so when this ratio is near unity x will be less than the mean. When c/C_u is in the neighborhood of one-half it will be necessary to try both possibilities.

These examples are intended only to illustrate the fact that distribution-free analyses can be accomplished when per unit understock costs are involved. Generally it is the case that inventory problems can be resolved on the basis of partial information about the demand distribution although the mathematics may become a bit sticky. We have given only some simpler examples which should serve to illustrate the general procedure.

26. COMPARISON OF ANALYSES WITH FULL AND PARTIAL INFORMATION

In the preceding two sections we have frequently given comparisons between the expected costs resulting from distribution-free analyses with those resulting if we assume that the demand distribution is normal. This was done purely for illustration and it can by no means be taken to be indicative of the general results of the two methods of analysis. The only generalization that we can certainly make is that any distribution-free analysis will always result in higher expected costs than will the same analysis if any specific demand distribution is assumed. If this were not the case we would be in the fortunate situation where information had no value and this is, to say the least, a rare event.

Our specific comparisons are not generalizable because they depend on too many specific features of the given inventory problem. To illustrate this fact let us return to the first such comparison we made, in Section 23. For our example we took $\bar{z} = 50$, $s = 5$, $x = 5$, $c = \$5$, and $K = \$1000$, a fixed cost for understock. The optimal order assuming normality was 61.75 with an expected cost of \$318. The optimal order under the Tchebycheff inequality was 71.55 with an expected total cost of \$412 or an actual expected cost of \$358 if the demand distribution was really normal without our knowing it. Obviously, any percentage comparison is strictly dependent on the size of $\bar{z}$. If $\bar{z}$ had been 100, other things remaining the same, the expected costs would have been \$568 for the normal distribution and \$662 for the distribution-free analysis. Thus, the percentage difference can be made as large or as small as desired simply by assuming a different $\bar{z}$. This suggests that we might confine our comparisons in this case to the amount of stock in

excess of $\bar{z}$ which is ordered under the two methods of analysis. Again on this basis it develops that the comparison may be quite good or it may be quite bad. That it may be quite bad is clear from the fact that the analysis based on the normal distribution may well result in the conclusion that no stock should be ordered, whereas the distribution-free analysis can never result in an order of less than $\bar{z}$. The reader can verify for himself that the comparison can be good by considering a case in which c/K is very small and s is also small.

Granted that no strictly correct generalizations can be made we will now proceed to make a generalization which we hope will be informative even though it is not strictly correct. This is that the distribution-free kinds of analysis will usually result in order sizes which have quite reasonable expected costs when compared to those resulting from analyses based on specific demand distributions. Remember that the only fair basis for comparison consists of comparing the actual expected cost of the distribution-free order size if the given distribution is assumed with the expected cost of the order size determined from assuming the given distribution. On this basis, we say, the distribution-free analyses compare surprisingly well. There are three main reasons for this and they deserve to be highlighted by enumeration.

First, there is a diminishing return, informationwise, from knowledge of higher moments of a distribution. When we say that a distribution is normal we are subsuming under this statement an infinite number of statements about the moments of the distribution. If knowledge of each successive moment would appreciably change our course of action we could hope to do very little indeed in the way of a distribution-free analysis. The facts are otherwise, however. The first four moments of a distribution alone carry almost all the information necessary to select a course of action. Thus, there can be, and is, success from distribution-free analysis.

Second, a marginal principle is at work on expected values and it results in a swapping of probabilities for dollars when we go to distribution-free analysis. This confusing statement can best be illustrated by example. Consider the example of Section 23 again. The optimal order under the assumption of normality would give a probability of being understocked of .94%. The distribution-free analysis under its own assumptions results in a probability of being understocked of 5.38%. If we had demanded the .94% level of protection from the distribution-free analysis we would have had to order 101.55 units, which would have been a pretty bad showing. This difference in protection levels is the result of the marginal principle we are referring to. Each one is, of course, optimal under its own assumptions. This effect is a basic reason why the distribution-free analyses do not do too badly.

Third, it is typically the case that the graph of expected cost against order size is quite flat in the central portion of its range. This means that fairly large divergences from the optimal order size may not change the expected cost by a comparable amount. We will have occasion to return to this point in a different context subsequently but we can offer some illustration of it here. For the example we have been using we have the following, under the assumption of normality:

ORDER SIZE	EXPECTED COST
50	\$750
52	605
54	482
56	395
58	345
60	323
62	318
64	323
66	331

For orders greater than 66 the expected cost simply goes up by \$5 for each additional item ordered. This table clearly shows the flat, slowchanging central portion of the curve of expected cost against order size. For this kind of situation it is more important to be in the right ball park than it is to know what team you are playing. The distribution-free analysis profits from this fact by being, generally, in the right ball park.

These, then, are the major reasons why distribution-free analyses do not do so badly as one might expect them to do. We are always better off if we have complete information but we are not often in this fortunate situation. The methods presented in this chapter should serve to demonstrate that there are sufficient resources available to permit us to utilize rationally such partial information as we may have.

PROBLEMS

1. Assume that in Problem 2 of the preceding chapter no information about the probability distribution of failures is available.
 - (a) What would be the minimax optimal strategy?
 - (b) What would be the optimal strategy by the regret criterion?
 - (c) What would be the optimal strategy by the Bayes criterion?
2. For the probability distribution of failures in Problem 2 of Chapter 2 calculate:
 - (a) The first five moments, μ_r .
 - (b) The first five moments around the mean, m_r .
 - (c) The absolute moments λ_1 , and λ_3 .
 - (d) The measure of skewness, γ_1 .
 - (e) The measure of kurtosis, γ_2 .
3. Assume that for a given item the per unit cost of overstock is \$10 and that the fixed understock cost is \$100.
 - (a) Given that the mean demand is 100 units and that the standard deviation is 10 units find the optimal amount of stock to order. Use both the Tchebycheff inequality and the more accurate expression illustrated at the end of Section 23.
 - (b) In addition to (a), it is known that the demand distribution is symmetrical. What is the optimal amount of stock to order?

- (c) Suppose it is known that $m_4 = 35,000$ and that the distribution is symmetrical. What is the optimal amount of stock to order?
 - (d) Suppose that it is known that $\gamma_2 = 0$. What is the optimal amount of stock to order?
 - (e) Suppose the information is as in (a) but that it is known that the distribution is unimodal with mode equal to 85. How much stock should be ordered?
4. Suppose that the per unit overstock cost for an item is \$10 and the per unit understock cost is \$40. Assume demand is symmetrically distributed with mean of 100 and standard deviation of 10. How much stock should be ordered?

dynamic models under certainty

27. GENERAL CHARACTERISTICS

The defining characteristic of dynamic inventory problems is that more than one order is possible. Dynamic models under certainty are ones which deal with dynamic inventory problems where the level of demand is known over the period of time involved. Demand may be at some constant rate over the given time period or it may vary but, in either case, it must be known in order to make the problem one under certainty.

These characteristics suggest some of the major structural features of the analyses of this kind of inventory problem. First, since demand is known with certainty there is generally no need to consider possibilities of overstock and understock. As a result, the two costs corresponding to overstock and understock will not enter into these analyses. Second, since multiple orders are possible we will need to introduce an order cost to reflect the penalties resulting from having more rather than less orders. Third, we will need to introduce a carrying cost to reflect the penalties resulting from maintaining a higher average level of stock over a period rather than a lower level. Our analyses will typically consist of a balancing of these two opposing costs.

Dynamic inventory problems under certainty are not frequently encountered. It is seldom that we know demand with certainty. However, when such problems do occur they can be important. The needs for materials for fabrication purposes or for production processes may well be known with certainty. For example, inventories of construction materials for a specific project must be maintained and the demand will frequently be known with certainty. A plant which is operating at a known level of production will require raw materials and the corresponding inventory problems are often

under certainty. Finally, many items for which there is actually a demand probability distribution nonetheless show such small variation in demand that their inventory problems can be effectively treated as being under certainty. As a matter of fact, we shall see that many practical dynamic inventory problems under risk can be most expeditiously handled by assuming that they are under certainty as a first approximation. Therefore, the topics discussed in this chapter can be of considerable usefulness.

28. OPTIMAL LOT-SIZE MODEL WITH CONSTANT DEMAND

Certainly the oldest and most widely known of all of the inventory models is that one which is called the optimal lot-size model. The reason for its relative antiquity is that the same basic derivation, and resulting equation, is useful in production planning for determining the optimal number of units to produce with one set-up of a machine or process. Industrial engineers were therefore the first ones to develop this particular analysis. It makes a good starting point for the analysis of inventory problems under certainty because of its simplicity.

In this kind of inventory problem we assume that demand for a particular item occurs at a constant, known rate over time. Whether we produce the item ourselves or order it from an outside producer there will be some cost associated with ordering the item. If we produce it ourselves we call it the set-up cost. If we order it from outside we call it the order cost. In either case, we will denote this cost by C_r (since we have already used C_o for the overstock cost). In addition, there is a carrying cost associated with carrying a unit of this item in stock for a given period of time. In the first chapter we discussed the components of this carrying cost and pointed out that it can usually be summarized most conveniently and usefully in the form of a percentage cost of a given number of dollars tied up in inventory for a given period of time. We will denote the carrying cost by C_c , which will always be expressed as a percentage per unit time.

Since these two costs are the only ones involved in our analysis we may well expect that they are opposing costs, in the sense discussed in Chapter 1. Of course, this is so but we must examine the relationship in detail. Let us introduce the required notation. The following symbols will be used:

- z = demand in a given period of time, in units,
- c = per unit cost of the item,
- C_r = cost per order,
- C_c = percentage cost per dollar in inventory for a given period of time,
- x = amount to order, in units.

We can choose any convenient period of time but both z and C_c should be expressed in terms of the same time unit. The cost per unit of the item, c , must include, of course, any assignable per unit costs involved in getting the item into inventory.

We need z units for the period of time. Now, we could get this number of units by a great variety of ordering patterns. At one extreme we could order all z units at the start of the period. At the other extreme we could order one unit at a time, thus requiring z orders. Obviously, the first extreme would minimize our ordering cost and the second extreme would maximize it. If this were the only cost there would be no problem—we would simply order once. We already know that there is an opposing cost which prevents this simple resolution. In this particular problem it is the carrying cost. Let us investigate how it behaves under our two extreme ordering policies. Figure 4-1 shows the typical saw-tooth pattern of inventory if we order x units on each order. Since the rate of demand is constant the amount in inventory, starting at x , decreases steadily until it reaches zero. If we are planning correctly—and there is no reason why we cannot since we know the constant rate of demand—our second order will arrive in stock just when the first order is completely depleted.

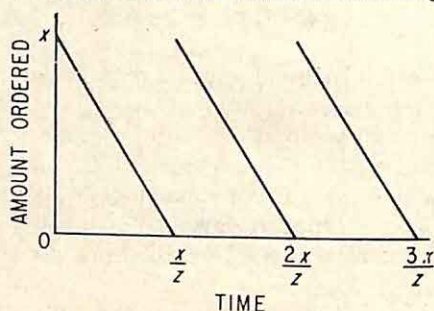


FIG. 4-1 Inventory behavior under certainty with constant rate of demand.

An application of elementary plane geometry is sufficient to demonstrate that, with such a pattern of inventory, the average amount in stock during the entire period will be simply one-half the amount ordered, or $x/2$. Thus, if we order all z units at once we will have a large average inventory throughout the period, namely $z/2$. If we order only one unit at a time we will have a very small inventory, namely, $\frac{1}{2}$ unit. In other words, our carrying costs decrease as our ordering costs increase, and vice versa. What we must do is to minimize the total costs which is equal to the sum of the order costs plus the carrying costs.

Fortunately, this is extremely easy to do. If we order x units at a time we will have to make z/x orders during the unit of time. This will cost $(z/x)C_r$. If we order x units at a time we will have an average inventory of $x/2$ units in stock constantly throughout the period. These units represent $(x/2)c$ dollars tied up in inventory and the carrying cost will be $(x/2)cC_c$. The total cost of any policy of ordering x units at a time will therefore be:

$$\text{Total Cost} = \text{T.C.} = \frac{zC_r}{x} + \frac{xC_c}{2}$$

We want to minimize this with regard to x so all we need do is differentiate with regard to x , set the derivative equal to zero, and solve the resulting equation. This gives the optimal order size:

$$x_o = \sqrt{\frac{2zC_r}{cC_c}}$$

Note that this equation gives the order size in units. If it is wanted in terms

of dollars one can calculate x and then multiply by c or simply do it in advance algebraically. The optimal order size in dollars is, then,

$$x_0 c = x_s = \sqrt{\frac{2czC_r}{C_c}}$$

Returning to the equation in units, let us illustrate its use on an item with a yearly demand of $z = 2000$ units, a cost of $c = \$3$, an order cost of $C_r = \$10$, and a carrying cost of $C_c = .12$. We may note that it is often the case that the order cost is higher than one would guess on the basis of uninformed "common sense." Putting these values into our equation gives $x_0 = 333$ units. This means that there should be six orders per year for this item. We could equally well have formulated this problem using the number of orders in a given time period, n , as the discretionary variable. Since we have already solved for x_0 we can determine the equation for n directly:

$$n = \frac{z}{x_0} = \sqrt{\frac{cC_c z}{2C_r}}$$

We might also have written the total cost equation in terms of the months between orders, assuming that the time period chosen was one year, as is frequently the case. In this event we would have minimized over this variable, $t =$ months between orders. We can construct the equation for the optimal value of t directly from the equation for n or x_0 :

$$t = \frac{12}{n} = \frac{12x_0}{z} = \sqrt{\frac{288C_r}{cC_c z}}$$

These last two equations may be more useful than the one for x_0 (naturally, they are all equivalent in terms of minimizing the total cost) because practical exigencies sometimes require that t be rounded off so that it is a divisor of twelve. By so doing it is ensured that the ordering cycle will repeat every year. How should we round off t if this is required? Our criterion must be to round off so that the total costs are kept as low as possible. The total cost equation written in terms of months between orders, t , is:

$$T.C_t = \frac{12C_r}{t} + \frac{tzcC_c}{24}$$

Now, of course, we know the optimal value of t and we could substitute it in this equation but let us forgo this for a moment. Instead, suppose we add some fraction of a month, d , to t and write the total cost equation for $(t + d)$ where d can be positive or negative:

$$T.C_{t+d} = \frac{12C_r}{t+d} + \frac{tzcC_c}{24} + \frac{dzcC_c}{24}$$

If t is the optimal value then $T.C_t < T.C_{t+d}$ so let us form the expression for the difference between the two total costs:

$$\begin{aligned} T.C_{t+d} - T.C_t &= 12C_r \left(\frac{1}{t+d} - \frac{1}{t} \right) + \frac{zcC_c d}{24} \\ &= 12C_r \left(\frac{1}{t+d} - \frac{1}{t} + \frac{zcC_c d}{288C_r} \right) \end{aligned}$$

But notice that

$$\frac{zcC_c d}{288C_r} = \frac{d}{t^2}$$

Therefore,

$$\begin{aligned} T.C_{t+d} - T.C_t &= 12C_r \left(\frac{1}{t+d} - \frac{1}{t} + \frac{d}{t^2} \right) \\ &= 12C_r \left[\frac{d^2}{t^2(t+d)} \right] \end{aligned}$$

This, then, is the expression which tells us how much the total costs will change if we round off the optimal value of t to some neighboring whole number. From the form of the expression it can be immediately seen that a change of a given positive amount will always be less costly than the same amount taken negatively.

We can use this expression to determine the optimal way of rounding off t to divisors of 12. Let us illustrate the argument. Suppose the optimal value of t falls between 2 and 3, say $t = 2 + k$. Then either $d = -k$, to round off to 2, or $d = 1 - k$, to round off to 3. When will the two ways of rounding be equal in their effects on total costs? From our equation for the effect of a change of d we can write:

$$\frac{k^2}{(2+k)^2(2)} = \frac{(1-k)^2}{(2+k)^2(3)}$$

Solving this gives $k = \sqrt{6} - 2 = .449$. Therefore, if $t \leq 2.449$, round off to 2 and if $k > 2.449$ round off to 3. Repeating this process we can prepare a table:

IF t IS BETWEEN	USE t EQUAL TO
0 - 1.414	1
1.415- 2.449	2
2.450- 3.464	3
3.465- 4.899	4
4.900- 8.485	6
8.486-14.697	12
14.698-20.785	18
20.789-	24

The last row indicates that it is acceptable to order only once every two years as a minimum—a company policy assumed for this particular case.

We have had several occasions to mention that it is typically the case that the cost curve, as a function of order size, is quite slowchanging in its central portion. This fact was helpful to us in the last chapter because it makes distribution-free analyses produce much better results than they would if the cost curve was rapidly changing. It can also be helpful in terms of the problems and difficulties of achieving good estimates of the relevant costs in an inventory problem. In the present kind of problem the only two costs involved are the order cost and the carrying cost and these are certainly easier to measure than is such a cost as the understock cost. However, as we discussed in the first chapter, even the order and carrying costs can be quite difficult to measure accurately. We will not repeat that discussion here. Instead, let us see how badly our solution may be affected by errors in the measurement of the two costs.

Substituting the optimal value of x in the total cost equation gives the optimal total cost:

$$\begin{aligned} T.C_{\text{opt}} &= \frac{zC_r}{\sqrt{\frac{2zC_r}{cC_c}}} + \frac{cC_c}{2} \sqrt{\frac{2zC_r}{cC_c}} \\ &= \sqrt{2cC_cC_rz} \end{aligned}$$

Now, let us assume that the two costs are actually C_c and C_r but that we think they are k_1C_c and k_2C_r . Using these values for the costs we would find the minimizing value of x :

$$x = \sqrt{\frac{2zk_2C_r}{ck_1C_c}}$$

The actual total cost for this erroneous value of x would be

$$T.C. = \left(\sqrt{\frac{k_1}{k_2}} + \sqrt{\frac{k_2}{k_1}} \right) \sqrt{\frac{zC_r c C_c}{2}}$$

The difference between this actual total cost and the true optimal total cost would be, expressed as a percentage of the optimal total cost

$$\frac{1}{2} \left(\sqrt{\frac{k_1}{k_2}} + \sqrt{\frac{k_2}{k_1}} \right) - 1$$

Using this expression we can now determine the effect of errors in our measurements of the two costs. The increase in total cost due to our misinformation about the order and carrying costs will be less than 10% of the actual minimum total cost as long as

$$.412 \leq \frac{k_2}{k_1} \leq 2.428$$

The extra total cost will be less than 6% of the actual minimum total cost if

$$.5 \leq \frac{k_2}{k_1} \leq 2$$

Thus, for example, suppose that we know C_r accurately; then as long as our estimate of C_c is between $\frac{1}{2}$ and 2 times its actual value, we know that the actual total cost will not exceed the true minimum total cost by more than 6%.

This shows that our analysis is not unduly sensitive to errors in the estimation of the relevant costs. This is fortunate since it means that the analysis can be useful even in cases where there is considerable latitude for error in the measurement of the costs. We will have the need to return to this subject again when we consider the design of an inventory study. The reader is advised to consult pages 176 through 181 for a further analysis of errors of estimation.

29. QUANTITY DISCOUNTS

A great many variants of this basic model can be constructed which take into account anticipated price changes, various kinds of limitations, quantity discounts, and so forth.* It will be sufficient for our present purpose to simply illustrate the general procedure involved for one kind of variant. We will consider the case where quantity discounts are offered by the vendor to the purchaser. In order to induce larger purchases the supplier often offers a reduced price if amounts greater than some minimum are ordered. In terms of our model this means that c becomes smaller if a large enough order is placed. The same effect may be produced because savings in transportation costs are available if a sufficient quantity of the item is purchased at one time. This also will result in a lower value of c for order amounts exceeding some given quantity. Our question is, should such a discount be taken or not?

There is clearly no problem in cases where the optimal x already exceeds the minimum amount necessary to obtain the discount. In this event the discount simply represents a bonus for an action which would have been taken anyway. The question we are considering arises when the optimal x is smaller than the amount necessary to obtain the discount. In this case we must determine whether the total costs which result from taking the discount will be less than those that result from ordering x . If so, we take the discount. If not, we will forgo the discount. There are three effects of a discount on total cost:

- (1) The reduction in c results in a saving per time period equal to the cost reduction per unit multiplied by the usage per time period.
- (2) Since the order size is larger if the discount is taken there are less orders per time unit and there is a saving in total procurement costs.
- (3) There is a change in carrying costs. The smaller c means that less inventory investment is required per unit but the larger number ordered results in more units in inventory.

Our procedure will be to take these three changes into account and mathe-

* For examples, see T. M. Whitin, *The Theory of Inventory Management* (Princeton, N.J.: Princeton University Press, 1953).

matically determine the circumstances under which a discount should be taken.

Assume that the optimal order size, x_0 , has been determined and that a discount of d is offered if kx_0 units are ordered, where k is greater than one. A discount of d means that the new cost of the unit becomes $(1 - d)c$. Let us now consider the three changes in total cost listed above:

(1) The direct saving in cost over the time unit will be zcd .

(2) The ordering cost if kx_0 are ordered each time will be zC_r/kx_0 .

(3) The carrying cost for an order size of kx_0 will be $\frac{1}{2}kx_0c(1 - d)C_c$. If it is to be advantageous to take the discount we must have the total cost of ordering kx_0 units minus the direct saving in cost less than the minimal total cost when x_0 units are ordered:

$$\frac{zC_r}{kx_0} + \frac{kx_0c(1 - d)C_c}{2} - zcd \leq \frac{zC_r}{x_0} + \frac{x_0cC_c}{2}$$

Substituting in the optimal value of x this becomes:

$$\frac{1}{k} \sqrt{\frac{zC_r c C_c}{2}} + k(1 - d) \sqrt{\frac{zC_r c C_c}{2}} - zcd \leq \sqrt{2zC_r c C_c}$$

or
$$\frac{1}{k} + k(1 - d) \leq d \sqrt{\frac{2cz}{C_r C_c}} + 2$$

For convenience, put $Q = \sqrt{\frac{2zc}{C_r C_c}}$. Then the equality holds when

$$k = \frac{2 + dQ \pm \sqrt{(2 + dQ)^2 - 4(1 - d)}}{2(1 - d)}$$

Since the quadratic equation involved has the second derivative with respect to k always positive, it follows that the equation has a minimum. Therefore, the values of k which fulfill the inequality are those lying between the two roots of the equality as given above. We want k to be as large as possible so it follows that the breakeven point for any discount, d , is given by

$$k = \frac{2 + dQ + \sqrt{(2 + dQ)^2 - 4(1 - d)}}{2(1 - d)}$$

For any given values of C_c and C_r tables can be prepared which show the breakeven value of k . Thus, for $C_c = .12$ per year and $C_r = \$10$ we have:

BREAKEVEN VALUE FOR k			
	d		
cz	.01	.02	.05
1000	1.901	2.458	3.991
5000	2.521	3.576	6.642
10000	2.986	4.446	8.780

The breakeven value of k is given in the body of the table. cz , of course, is

simply the dollar demand value for the time period in question, here one year. If less than kx_0 units will get the discount it will be advantageous to order the least amount necessary in order to get it. The farther the minimum amount necessary to get the discount is below kx_0 , the greater the gain from the discount. If the amount which must be ordered to obtain the discount is greater than kx_0 then the discount should not be taken and the usual x_0 units should be ordered.

30. OPTIMAL LOT-SIZE MODEL WITH VARYING DEMAND*

When the demand is known with certainty it may still be the case that it is not at a constant rate. For example, seasonal fluctuations in production level may require that some raw material, of which the amount needed is a known function of the known production level, should be available in specified, varying amounts. We would like a method for determining the optimal ordering policy for such a case as this. As an example of such a case let us assume that we have the following known amounts of a particular item which will be needed:

MONTH	DEMAND
1	40
2	70
3	90
4	60
5	50
6	30
7	20
8	20
9	50
10	70
11	100
12	60

We will assume that this demand pattern is repeated year after year. Let us suppose, further, that the ordering cost, C_r , is \$10, the carrying cost, C_c , is 1% per month, and that for the item in question the cost $c = \$4$ per unit. The problem is to determine the optimal ordering policy for this item.

As a general preliminary we may note that it is always possible to write some kind of mathematical expression which represents a given problem and then to state that the solution of the problem consists in performing some kind of manipulation of the expression. In the present case, for example, we could write an expression which would correctly represent this particular problem and then state that the solution to the problem consisted in minimizing the expression. However, it is often the case that such a procedure,

* For a more general statement of this model see Harvey M. Wagner and Thomson M. Whitin, "Dynamic Version of the Economic Lot Size Model," *Management Science*, 5, 1 (October, 1958), pp. 89-96.

while true in theory, is of no advantage whatsoever in practice. This results whenever it is not feasible to actually perform the stated manipulation of the mathematical expression. Such would be the case in the present instance. It is of no avail to write the equation specifically representing this problem because it is not one which we could minimize, at least with any reasonable amount of work. Under such circumstances we try to proceed in an entirely different way. Instead of looking for an equation which we can solve we look for an algorithm, as it is called. An algorithm is a procedure which, when followed, will lead to the solution of the given problem by an iterative process. Compared to solving solvable equations, an algorithm is not so "neat" and it requires more computation. However, compared to solving complex equations for which an algorithm is a reasonable substitute, the algorithm is a pleasure and a delight. It is an algorithm which we need in the present case. The one which we will use is called dynamic programming and is applicable to a most remarkable range of problems apart from the present one.

We need to make some assumption about the way the carrying cost is to be calculated. Sufficiently realistic for our purposes, and relatively simpler, will be the assumption that the carrying cost is only applicable to inventory which is carried over from one month to the next. In other words, we will not credit any carrying cost in a given month to the stock which is used in that month. Thus, suppose we order, in the first month 110 units. At the end of the first month we would calculate the carrying cost for the 70 units carried over to the second month but not for the 40 units used in the first month. We will further assume that the delivery time after ordering is known with certainty and that all orders will be placed so that they are delivered on the first day of the month. This is not a particularly restrictive assumption because, if it is unsatisfactory, we need only present our data in a weekly form, for example, and follow the same procedure which we are going to develop.

The dynamic programming method which we are going to use does not sound very grandiloquent when it is summarized in words since it consists of a procedure for trying all possible ordering policies and picking that one which is best. However, the essence of the method is a logical procedure by means of which we have only actually to calculate a small subset of all possible policies in order to find the best one. It is the enormous saving in calculation which results from dynamic programming that makes this algorithm a feasible one. Instead of presenting a mathematical description of the algorithm we will reason our way through the present problem and develop the dynamic programming logic as we go.

The alternative policies which we must consider result from the obvious fact that for each month we have two alternatives. We can either order that month's demand for delivery at the start of the month or we can have had it already in stock from the preceding month. Note that there is no need to consider any intermediate possibility. It would never be to our advantage to carry some stock forward to meet the month's demand and to simultaneously order the rest. This is obviously true since we could instead order

the total amount for the month with no increase in ordering cost and with a saving of the carrying cost for the part of the demand carried forward. Thus, we need consider only two possibilities for each month: order the demand for the month at the start of the month or carry forward inventory to meet the total demand of the month.

A second, and very important, restriction is equally obvious. This is that there is an upper limit to the number of months in advance it could possibly be to our advantage to carry stock forward. Consider the demand in the eleventh month as an example. If we ordered stock to be delivered at the start of the eighth month to meet this demand we would have to carry it for three months. The cost of carrying 100 units for three months would be

$$100(3)(\$4)(.01) = \$12$$

This is greater than the order cost so it would clearly be to our advantage to order the 100 units at the start of the eleventh month rather than to carry them from the eighth month. In other words, it would never be to our advantage to carry stock for the eleventh month from the eighth, or any earlier, month. Alternatively, we can say that the planning horizon for the eighth month cannot exceed the tenth month. Every month has some such limitation to its planning horizon. Our first step will be to determine what these limitations are. We can do this conveniently in a tabular form which will be simultaneously of great usefulness in our subsequent calculations. We will prepare a table which will show the total carrying costs for each month assuming that we order in that month for n months in advance. We will do this for each month up to the limit of that month's planning horizon. It is easy to do this by a systematic procedure. We will present the table first and then explain how we got it:

MONTHS ORDERED	UNITS CARRIED											
	MONTH											
	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	70	90	60	50	30	20	20	50	70	100	60	40
3	250	210	160	110	70	60	120	190	270	220	140	180
4	430	360	250	170	130	210	320	490	450	340	350	450
5	630	480	330	250	330	490	730		610	620	710	
6	780	580	430	500	680				960			
7	900	700	730	920								
8	1040	1050										
9	1440											

The first row of this table is all zeros since according to our assumption on carrying costs, there is no carrying charge if stock is not carried forward. Thus, since the first row shows the units carried if one month's demand is ordered in each of the months, there are no units carried forward and, hence, no carrying charges. We can use units carried as a direct measure of carrying charges because the carrying charges are simply $\$4(.01)$ times the number of units carried. The second row shows the number of units that will be carried

if two month's demand are ordered in each of the months. Thus, this row is simply the monthly demands displaced one to the left. If we order two month's demand in the first month we will have to carry the demand for the second month, 70, for one month. This is the entry shown. Now let us proceed by columns and take the first month's column as an example. If we order three month's demand in the first month we will carry 70 units for the second month's demand for one month plus 90 units for the third month's demand for two months or $70 + 2(90) = 250$ units for one month. For four months ordered we simply add to this 60 units for the fourth month carried for three months or $250 + 3(60) = 430$ units for one month. Continuing in this way we can quickly fill out the table. When do we stop in any column? The order cost is \$10 and this is equivalent to carrying 250 units for one month since $250(\$4)(.01) = \$10$. Therefore, as soon as the difference between two successive entries in a column equals or exceeds 250 we can stop at the smaller entry because we know that we could do as well or better by adding another order. Thus, consider the column for the sixth month. If we order four month's demand we have to carry the equivalent of 210 units for one month. If we order five month's demand we have to carry the equivalent of 490 units for one month, an increase of 280 units or an increased cost of \$11.20. This increase is due to having to carry the tenth month's demand, 70, for four months. Clearly, we can improve this policy by simply ordering the tenth month's demand at the start of the tenth month, at a cost of \$10 and a saving of \$1.20, similarly for every month. The underlined entry in each column is the limit of the planning horizon for that month, determined in this way.

Now, if this problem were for one year only we could start our programming procedure with the knowledge that there had to be an order in the first month. But in our case we are endeavoring to determine the optimal order policy under the assumption that the same pattern of demand will be repeated for a number of years. We do not know, therefore, any specific month in which there will definitely be an order. As a result, we will have to try several possibilities since the procedure requires that a month be known in which there will definitely be an order. To economize on our efforts we want to select a starting point with the minimum possible number of such alternatives. The month with the shortest planning horizon provides us with this minimum number of alternatives. We have two months with a planning horizon of three months. Take the eighth month, one of the two, as an example. We know that it can never be to our advantage to order the eleventh month's demand in the eighth month. Therefore, since it automatically follows that it would be even more to our disadvantage to order the eleventh month's demand in any month preceding the eighth, we must order the eleventh month's demand in either the ninth, the tenth, or the eleventh month. Our procedure, then, will be to assume successively that there is an order in the ninth, the tenth, and the eleventh months. For each of these three alternatives we will determine the minimum cost ordering policy. The over-all minimum from among these three will, then, be the optimal ordering policy for our problem.

We will only go through the argument in detail for the first alternative—assuming that there is an order in the ninth month. The other two alternatives are simply repetitions of the same reasoning so nothing would be gained by repeating the argument for them. We start, then, by assuming that an order is placed in the ninth month. We want to minimize the total ordering and carrying costs associated with an ordering policy. Since we have converted carrying charges to an equivalent units carried in one month we can most easily do likewise with the order cost. The \$10 order cost is equivalent to the carrying charges for 250 units for one month. We will use this equivalent and make all our calculations in terms of equivalent units carried for one month. We can convert into dollars at any point, if so desired, by multiplying by .04. Since we are assuming that there will be an order in the ninth month it follows that the optimal policy for the ninth month's demand is to order it in the ninth month. This has an order cost of 250 and no carrying cost. We will write this optimal policy as (9). The essence of the dynamic programming method is to proceed seriatim, assuming at every step that we have determined the optimal policy through the preceding step and then determining the optimal policy for the next step on that basis. The idea will become clearer as we proceed. Let us consider the policy for the tenth month. There are two alternatives: order enough in the ninth month for the tenth month, designated (9-10), or make a separate order for the tenth month, designated (10;9). The order cost for (9-10) is 250 and we can look in our table to determine the carrying cost, 70, or a total cost of 320. For (10;9) we have two orders and no carrying cost or a total cost of 500. However, as we shall see, what we are really doing here is to add the cost of the policy described before the semicolon in (10;9) to the cost of the optimal policy coming after the semicolon. In any event, of the two alternatives it is evident that (9-10) has the lower cost. It, therefore, is the optimal policy for the tenth month, assuming the optimal policy for the ninth month.

We now proceed to the eleventh month. There are three alternatives: order for the eleventh month in the ninth month, (9-11), order for the eleventh month in the tenth month and use the optimal policy for the ninth month, (10-11;9), or order in the eleventh month and use the optimal policy for the tenth month, (11;10). We look up (9-11) directly in our table—remembering to add the 250 for ordering cost—and find 520. We calculate (10-11;9) by looking 10-11 up in the table, 100, adding 250 for the order in 10, and adding the cost of the optimal policy for the ninth month, 250. This is a total cost of 600. (11;10) is simply 250 for the order in 11 plus the optimal policy cost for 10, 320, or a total of 570. The lowest cost for the eleventh month, and the optimal policy for this month, is (9-11) with a cost of 520.

The four alternatives for the twelfth month are (9-12), (10-12;9), (11-12;10), and (12;11). As an example of these calculations consider (11-12;10). We look up 11-12 in our table and find 60, add 250 for the order in 11, and add the 320 cost of the optimal policy for 10. Thus, the total cost for (11-12;10) is 630. This happens to be the minimum cost and the optimal policy for the twelfth month. The advantages of this method of calculation should be clear. When we want to calculate the cost of ordering the twelfth

month's demand in the eleventh month we do not need to consider all of the possible ways of ordering the ninth and tenth months' demands because we have already determined the minimum cost way of ordering these two months' demands. The farther we go in our calculations, the greater the advantage of this saving in the number of possibilities that must be considered. We can routinize the calculations in a table as follows, where we have simplified the notation in an unambiguous way by letting our previous symbol $[(t+1)-(t+k);t]$, be represented by the equivalent $(t+k;t)$:

MONTH					
9	10	11	12	1	2
<u>(9) 250</u>	<u>(9-10) 320</u>	<u>(9-11) 520</u>	<u>(9-12) 700</u>	<u>(9-1) 860</u>	<u>(2;10) 920</u>
	(10;9) 500	(11;9) 600	(12;9) 720	(1;9) 840	(2;11) 950
		(11;10) 570	(12;10) 630	(1;10) 710	(2;12) 950
			<u>(12;11) 770</u>	<u>(1;11) 810</u>	<u>(2;1) 960</u>
				(1;12) 880	

MONTH					
3	4	5	6	7	8
<u>(3;12) 1130</u>	<u>(4;12) 1310</u>	<u>(5;12) 1510</u>	<u>(6;12) 1660</u>	<u>(7;12) 1780</u>	<u>(8;12) 1920</u>
(3;1) 1050	(4;1) 1170	(5;1) 1320	(6;1) 1440	(7;1) 1540	(8;1) 1660
(3;2) 1170	(4;2) 1230	(5;2) 1330	(6;2) 1420	(7;2) 1500	(8;2) 1600
	(4;3) 1300	(5;3) 1350	(6;3) 1410	(7;3) 1470	(8;3) 1550
		(5;4) 1420	(6;4) 1450	(7;4) 1490	(8;4) 1550
			(6;5) 1570	(7;5) 1590	(8;5) 1630
				(7;6) 1660	(8;6) 1680
					(8;7) 1720

(Thus, for example, we read 5;2 as: (a) order in third month for the third, fourth and fifth months, (b) use optimal policy of the second month, namely, 2;10. We read 2;10 as: (c) order in the eleventh month for the eleventh, twelfth, first and second months, (d) use optimal policy of the tenth month, namely 9-10. We read 9-10 as: (e) order in the ninth month for the ninth and tenth months.)

The lowest cost, hence optimal policy, for each month is underlined. Why are some of the middle columns shorter than their neighbors? This is the result of our other method of eliminating unnecessary alternatives—the limited planning horizons of the various months. For example, consider the column for the second month. Two alternatives are missing: (9-2) and (2;9). These are missing because the planning horizons of the ninth and the tenth months are five months and four months, respectively. Neither of these therefore includes the second month so they need not be calculated since we know they will be more costly than other alternatives.

Verbal descriptions of algorithms are always more lengthy than the process itself. Take the algorithm for calculating the square root of a number as an example. In words it is interminable and sounds very complex. Yet it is actually a most simple and expeditious way for finding square roots. So it is with this dynamic programming algorithm. Only by trying one will the reader convince himself that it is actually a simple and rapid procedure. The

computations required up to this point can easily be accomplished in half an hour.

In any event, we are through with this part of the analysis. What have we accomplished? We have found the minimum cost ordering policy under the assumption that an order is placed in the ninth month. The total cost is given by the underlined figures in the last column, 1550 or, in dollars, $1550(.04) = \$62$. What is the policy? We have only to read the columns backwards. From column 8 we see that we can either order in the fourth or fifth months (since there is a tie) for all the months from 4 (or 5) through 8 inclusive. We will take the first of these. Specifically, our instruction reads: Order for the fourth, fifth, sixth, seventh and eighth months in the fourth month and use the optimal policy for the third month. Looking at the third month's column we see that we must order the second and third months' demands in the second month and use the optimal policy for the first month. Looking in the column for the first month we see that we must order for the eleventh through the first months in the eleventh month and use the optimal policy for the tenth month. Looking in the column for the tenth month we see that we must order the ninth and tenth months' demands in the ninth month—and this finishes it. Let us summarize the optimal policy and verify the costs:

ORDER IN	ORDER FOR	ORDER COST	CARRYING COST
4	4-8	250	250
2	2-3	250	90
11	11-1	250	140
9	9-10	250	70
		<u>1000</u>	<u>550</u>

It can be seen that the costs check out to 1550, as they should. The reader can verify that the alternative minimum cost ordering policy gives the same total cost.

This, then, is the optimal ordering policy if an order must be placed in the ninth month. We have still to check what the optimal ordering policies will be when we require an order in the tenth or the eleventh month. We will simply give the results of the calculations for these two cases and leave it to the reader to verify them. For an order placed in the tenth month the optimal order policy is:

ORDER IN	ORDER FOR	ORDER COST	CARRYING COST
5	5-9	250	330
2	2-4	250	210
10	10-1	250	340
		<u>750</u>	<u>880</u>

The total cost is 1630 in units or \$65.20. For an order placed in the eleventh month the optimal order policy is:

ORDER IN	ORDER FOR	ORDER COST	CARRYING COST
9	9-10	250	70
4	4-8	250	250
2	2-3	250	90
11	11-1	250	140
		<u>1000</u>	<u>550</u>

Thus, the total cost if the order must be placed in the eleventh month is 1550, the same as we found when the order had to be placed in the ninth month. Therefore, either of these ordering policies costs \$62 and either one of them can be used.

Thus, we have solved this problem. How do we know we have the optimal ordering policy? We know that there has to be an order either in the ninth or the tenth or the eleventh month. For each of these three possibilities we have determined the optimal policy. We know that the policies so determined are, in fact, the optimal ones because our method of calculation is equivalent to trying out all possibilities and picking the best one. All we have done is to arrange our procedure so we could eliminate obviously inefficient policies without calculating their costs. Therefore, when we select the minimum cost policy among our three optimums we know that we have found the optimal ordering policy for the problem as originally stated.

31. MORE THAN ONE ITEM: THE OPTIMAL POLICY CURVE

Let us return to a consideration of the basic optimal lot-size model but extended to the case where there are a number of items for which inventories must be maintained. This, of course, is the more typical case. Very few organizations have only one item which they must maintain in inventory and, as organizations get larger, the trend is very decidedly in the direction of enormous numbers of items which must be maintained in stock. Some organizations number their stock-keeping units in the hundreds of thousands. We want to see whether the existence of a multitude of items changes the problems, adds new problems, or increases the range of possible solutions to the problems.

On first view there appears to be no reason to expect any changes simply because of the existence of large numbers of items in inventory. The equation we derived for the optimal amount to order contained no term representing number of items in inventory nor did this question arise in any way in our argument. Further, it is evident that the same equation can be applied to each item, using the data for that item, and that the optimal amount to order for the item will be determined by the equation. The sum of the optimal policies for each item will represent the best over-all solution to the total inventory problem. Therefore, there is no reason to expect any change in our analysis resulting from the hypothesized large number of items.

All of these remarks are true but, first, there are some new problems which

can arise—we will consider one such in the next section and a minor example in this section—and, second, there are some new possibilities which arise for the resolution of the inventory problems. It is this which we shall be mainly considering in this section. (In Part 2 we consider other methods for the resolution of many item problems. The reader is advised to consult the Index.)

The new possibilities arise in connection with the problem of measuring the costs required by the optimal lot-size equation. We tried to emphasize in the first chapter that there can be some serious difficulties in the way of successfully estimating these costs. The worst of these difficulties stems from the fact that we really must measure opportunity costs and cost accounting data is not typically suitable for this purpose. However, we will not repeat the discussion of the first chapter concerning these questions. Instead, we will simply develop some means for avoiding these problems which exist in the case of inventories of many items.

We will need an illustrative example. Suppose that a company maintains inventories of five items. The yearly demand, z_i , and the cost per unit, c_i , for these items are:

ITEM	z_i	c_i
1	600	\$3
2	900	10
3	2400	5
4	12000	5
5	18000	1

Typically, if one were called upon to make a study of this company's inventory problem one would find that the company was using some ordering system such as ordering each item once a month. Assuming that this is the case, we can calculate the number of orders per year and the average inventory which results:

ITEM	ORDERS PER YEAR	AVERAGE INVENTORY
1	12	\$75
2	12	375
3	12	500
4	12	2500
5	12	750
	60	\$4200

where the average inventory is easily calculated from $(z_i c_i)/24$.

If, now, we are told that the ordering cost for this company is \$10 per order and the carrying cost is 12% per year, we can calculate the total cost of the company's current ordering policy:

$$T.C_{\text{Current}} = 60(10) + .12(4200) = \$1104$$

Knowing that $C_r = \$10$ and that $C_c = .12$, we can easily calculate the optimal

ordering policy for each of the items. We will use the expression for order size in terms of dollars rather than in terms of units:

$$x_{is} = \sqrt{\frac{2c_i z_i C_r}{C_c}}$$

Using this expression we can quickly calculate:

ITEM	OPTIMAL ORDER SIZE IN DOLLARS	ORDERS PER YEAR
1	\$548	3.28
2	1225	7.35
3	1414	8.49
4	3162	18.98
5	1732	10.39
	\$8081	48.49

where orders per year are calculated directly from $(z_i c_i)/x_{is}$. Since the average inventory for each item is simply one-half the order size in dollars we can immediately determine the total cost of the ordering policy we have calculated:

$$T.C_{opt} = 10(48.49) + .12(4040) = \$970$$

We will not consider here the question of rounding off the order periods but will assume that the company can equally well use any desired order period for an item. It will be seen that the utilization of our equation for the optimal policy has resulted in a decrease in total costs of \$134, or about 12%.

Before we continue the intended direct line of discussion let us take a sidetrack and consider a kind of problem which might arise but which would not require any analysis in the case of one item. Suppose that the company has a shortage of working capital and cannot afford to maintain the indicated optimal average inventory investment of \$4040. Precisely, let us suppose that the most the company can afford in inventory investment is an average \$3000. How should the ordering policy be changed to minimize the total costs under this restriction? Let us formulate the problem mathematically. Using our summation notation, the total cost equation when x_i units of the i th item are ordered is:

$$T.C. = 10 \sum_i \frac{z_i}{x_i} + .12 \sum_i \frac{x_i c_i}{2}$$

We have been given the restriction that

$$\sum_i \frac{x_i c_i}{2} = 3000$$

Our problem is to minimize T.C. subject to the given restriction. The mathematical technique for accomplishing this is to use Lagrangian multipliers. This rather formidable sounding technique is actually very easy to use. We write the restriction as an equation equal to zero:

$$\sum_i \frac{x_i c_i}{2} - 3000 = 0$$

Now we form the Lagrangian, denoted L , by adding λ , the multiplier, times this equation to the expression to be minimized, T.C. Thus, we have

$$L = 10 \sum_i \frac{z_i}{x_i} + .12 \sum_i \frac{x_i c_i}{2} + \lambda \left(\sum_i \frac{x_i c_i}{2} - 3000 \right)$$

To minimize the original expression subject to the restriction we must minimize L over the x_i and λ . We do this by the usual procedure of taking derivatives and setting them equal to zero and we get:

$$\frac{\partial L}{\partial x_i} = -\frac{10z_i}{x_i^2} + .06c_i + \frac{\lambda c_i}{2} = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_i \frac{x_i c_i}{2} - 3000 = 0$$

These simultaneous equations can be readily solved by determining x_i from the first equation:

$$x_i = \sqrt{\frac{20z_i}{(.12 + \lambda)c_i}}$$

substituting this in the second equation:

$$\sqrt{\frac{20}{.12 + \lambda}} \sum_i \sqrt{z_i c_i} = 6000$$

and solving for λ :

$$\lambda = 20 \left(\frac{\sum_i \sqrt{z_i c_i}}{6000} \right)^2 - .12$$

In order to obtain a numerical value for λ we must determine the square of the sum of the square roots of $z_i c_i$ from the original data. We have

ITEM	$\sqrt{z_i c_i}$
1	42.43
2	94.87
3	109.54
4	244.95
5	134.16
	625.95

From this we immediately calculate

$$\lambda = .09767$$

Now we can determine the x_i from the equation above which includes λ . Even more conveniently we can calculate

$$\begin{aligned} x_{i,s} &= \sqrt{\frac{20}{.12 + \lambda}} \sqrt{z_i c_i} \\ &= 9.585 \sqrt{z_i c_i} \end{aligned}$$

since we have already determined $\sqrt{z_i c_i}$. Thus, we obtain

ITEM	ORDER SIZE IN DOLLARS	ORDERS PER YEAR
1	\$406.70	4.43
2	909.30	9.90
3	1049.90	11.43
4	2348.00	25.56
5	1286.00	14.00
	\$5999.90	65.32

The total cost for this ordering policy is

$$\text{T.C.} = 10(65.32) + .12(3000) = \$1013$$

This is the minimum cost ordering policy of all possible policies which would give an average inventory investment of \$3000. It will be seen that the total cost has gone up as compared to the optimal policy—as it obviously must. However, for an increase in total costs of only \$43 the company has accomplished its purpose of cutting its average inventory investment by \$1040.

Our sidetrack has fulfilled the subsidiary purpose of introducing the extremely useful device of Lagrangian multipliers as a way to minimize expressions subject to restrictions. We will need this device again when we return to our main theme. Suppose, now, that there are grievous difficulties in the way of measuring the two costs necessary to determine the optimal ordering policy for the company. Our analysis of the optimal ordering policy subject to the requirement that the average dollar inventory should not exceed \$3000 suggests an approach which might be used. The existing company policy produces an average inventory of \$4200 and uses 60 orders per year. Now, with no knowledge of C_r or C_c whatsoever, we can certainly reason that if we can reduce inventory and maintain the same number of orders per year then we have improved the company's position. This must be true, regardless of the costs involved. We propose, then, to see if we can decrease the average inventory investment while keeping the number of orders per year fixed at 60. If we can do this we will have improved the situation even though we will not be able to determine the actual amount of the saving.

We will proceed exactly as before. We want to minimize the inventory subject to the restriction that total orders per year will be 60. The expression for total average inventory is:

$$\text{Total average inventory} = \text{T.I.} = \sum_i \frac{x_i c_i}{2}$$

We want

$$\text{Total orders} = \text{T.O.} = \sum_i \frac{z_i}{x_i} = 60$$

We form the Lagrangian:

$$L = \sum_i \frac{x_i c_i}{2} + \lambda \left(\sum_i \frac{z_i}{x_i} - 60 \right)$$

and set the derivatives equal to zero:

$$\frac{\partial L}{\partial x_i} = \frac{c_i}{2} - \frac{\lambda z_i}{x_i^2} = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_i \frac{z_i}{x_i} - 60 = 0$$

Solving the first equation for x_i gives

$$x_i = \sqrt{2\lambda} \sqrt{\frac{z_i}{c_i}}$$

and substituting this in the second equation and solving for λ gives

$$\lambda = \frac{\left(\sum_i \sqrt{z_i c_i} \right)^2}{7200}$$

We already have determined $\sum_i \sqrt{z_i c_i}$ for these data so we calculate directly

$$\lambda = 54.42$$

From this we now obtain

$$x_{is} = x_i c_i = \sqrt{2\lambda} \sqrt{z_i c_i} = 10.43 \sqrt{z_i c_i}$$

This gives

ITEM	ORDER SIZE IN DOLLARS	ORDERS PER YEAR
1	\$442.70	4.07
2	989.80	9.09
3	1142.80	10.50
4	2554.90	23.48
5	1399.30	12.86
	\$6529.50	60.00

The inventory is one-half of the order size in dollars or \$3265. The number of orders has remained at 60, our restriction, and the inventory has been reduced by \$935, or 22.3% of its amount under the current policy. This is the minimum possible inventory investment with 60 orders per year. The decrease in inventory investment has been achieved by redistributing the orders among the items in an optimal way. We cannot, of course, determine what this reduction in inventory is worth to the company because we do not know the value of C_c .

The success of this analysis immediately suggests the corresponding argument in reverse. Why not improve the company's position, without

knowledge of the costs, by keeping the average inventory investment at the same level and decreasing the total number of orders per year? This can also be done. The expressions for T.I. and T.O. are the same as before, but their roles are reversed. We want to minimize

$$\text{T.O.} = \sum_i \frac{z_i}{x_i}$$

subject to

$$\text{T.I.} = \sum_i \frac{x_i c_i}{2} = \$4200$$

We form the Lagrangian:

$$L = \sum_i \frac{z_i}{x_i} + \lambda \left(\sum_i \frac{x_i c_i}{2} - 4200 \right)$$

and set the derivatives equal to zero

$$\frac{\partial L}{\partial x_i} = -\frac{z_i}{x_i^2} + \frac{\lambda c_i}{2} = 0$$

$$\frac{\partial L}{\partial \lambda} = \sum_i \frac{x_i c_i}{2} - 4200 = 0$$

We find

$$x_i = \sqrt{\frac{2z_i}{\lambda c_i}}$$

$$\lambda = \frac{\left(\sum_i \sqrt{z_i c_i} \right)^2}{35,280,000}$$

For our data

$$\lambda = .01111$$

and

$$x_{is} = 13.42 \sqrt{z_i c_i}$$

Therefore, we can calculate:

ITEM	ORDER SIZE IN DOLLARS	ORDERS PER YEAR
1	\$569.40	3.16
2	1273.10	7.07
3	1469.90	8.16
4	3286.40	18.26
5	1800.00	10.00
	\$8398.80	46.65

Average inventory remains at \$4200, within rounding errors, while total orders per year have been reduced to 46.65, a saving of 13.35 orders per year or 22.3% of the orders under the current policy. This policy is the one which has the minimum total orders per year of all policies which would give an

average inventory investment of \$4200. It will be noted that the percentage saving is the same in both cases, 22.3%. This will always be the case. The question of which of these two possible improved policies is preferable cannot be answered without knowledge of the relevant costs. However, either of them will clearly save the company money—even though we cannot calculate how much.

Now we have calculated three different optimal policies under different restrictions and we calculated the original optimal policy for assumed values of C_r and C_c . Careful inspection of these four different optimal policies indicates an underlying uniformity among them. This is shown in the following table:

POLICY	TOTAL INVENTORY	ORDERS PER YEAR	PRODUCT (T.I.)(T.O.)
Optimal; $C_r = \$10$, $C_c = .12$	\$4040	48.49	195900
Optimal; Inventory = \$3000	3000	65.32	195960
Optimal; Orders = 60	3265	60	195900
Optimal; Inventory = \$4200	4200	46.65	195930

All the products of T.I. times T.O. are equal to a constant value, within rounding errors. Further, with a few tries it is possible to discover where this constant value, 195,900, originates. More precisely, the constant value is 195,910 and it is equal to

$$\frac{1}{2}(625.95)^2 = \frac{1}{2}\left(\sum_i \sqrt{z_i c_i}\right)^2$$

We must certainly undertake to explain the reason for this structural similarity among the various optimal policies.

Reviewing the equations for x_{is} which we used in determining the actual optimal ordering policies in each of the four cases we see that in each of them we had

$$x_{is} = K\sqrt{z_i c_i}$$

K , of course, differed in the four cases but it was always determined either from the costs, C_r and C_c , or from λ . Presumably this is the reason for the occurrence of $(\sum \sqrt{z_i c_i})^2$ in the constant product of T.I. and T.O. However, instead of pursuing this line of reasoning we will take a different route which will lead us to the same goal. Let us write the expression for T.I. in terms of the optimal x_{is} , as determined in the optimal lot-size model:

$$\text{T.I.} = \sum_i \frac{x_{is}}{2} = \sum_i \frac{1}{2} \sqrt{\frac{2c_i z_i C_r}{C_c}} = \sqrt{\frac{C_r}{C_c}} \sum_i \sqrt{\frac{c_i z_i}{2}}$$

Similarly, the expression for T.O. is

$$\text{T.O.} = \sum_i \frac{z_i c_i}{x_{is}} = \sqrt{\frac{C_c}{C_r}} \sum_i \sqrt{\frac{c_i z_i}{2}}$$

These two expressions are valid for any values of C_r and C_c whatsoever—plugging in any given values and the data on $z_i c_i$ will give the optimal total inventories and total orders per year. Now, the interesting thing about these two expressions is that when they are multiplied together the two costs, C_r and C_c , cancel out and we are left with

$$(T.I.)(T.O.) = \frac{1}{2} \left(\sum_i \sqrt{z_i c_i} \right)^2$$

exactly the constant product we already discovered empirically from our four optimal policies. Thus, we see that the product of T.I. and T.O., for the optimal policy based on any possible values of the two costs, will always be a constant for the given data. Further, we see that if we divide our two expressions everything cancels out except the two costs:

$$\frac{T.I.}{T.O.} = \frac{C_r}{C_c}$$

Therefore, given the T.I. and T.O. for any optimal policy we can immediately determine the ratio of the two costs upon which the optimal policy was constructed.

This last fact suggests an alternative interpretation of the optimal policies we determined under various restrictions. When we restricted the inventory to \$3000 we found the optimal policy had 65.32 orders per year and, of course, \$3000 in inventory. The ratio, $(T.I.)/(T.O.)$, is, therefore, for this policy equal to 45.9. If the order cost is known to be \$10 we can conclude that this policy is imputing $C_c = \$10/45.9 = .218$. In other words, the policy we determined would only be optimal if $C_c = .218$. Therefore, the imputed value of C_c must be .218. Now, this policy resulted because the company had a shortage of working capital and wanted to restrict their inventory investment to \$3000. We have discovered that this restriction automatically imputes a value of .218 to the carrying cost. We can reason similarly for the other two optimal policies under restrictions. Restricting orders to 60 and assuming, for example, that the carrying cost is .12 automatically imputes a value of \$6.53 to C_r . Restricting inventory to \$4200 and assuming that the order cost is \$10 automatically imputes a value of .111 to the carrying cost.

The relation between these imputations and the various λ 's we calculated is an interesting one. For our first restriction we assumed that we knew that $C_r = \$10$ and $C_c = .12$ but that the company was forced to restrict the inventory to \$3000. We found $\lambda = .09767$ and we note that $.12 + .0977 = .218$, the imputed carrying cost in this case. In other words, λ is not an arbitrary mathematical device. Its value is equal to the increment added to the carrying cost because of the inventory restriction. It plays, in short, the role of a cost in this formulation. The second restriction we imposed was that orders should equal 60. We found $\lambda = 54.42$ and it will be noted that $54.42(.12) = \$6.53$, the imputed value of C_r for $C_c = .12$. Therefore, λ in this case is the ratio C_r/C_c . The third restriction was that inventory should

equal \$4200. We found $\lambda = .01111$ and $10(.01111) = .111$ is exactly the imputed carrying cost for this case. Here, in other words, λ is the ratio C_c/C_r . In every case, λ is intimately related to the imputed costs resulting from restrictions. In most economic problems in which it is necessary to use Lagrangian multiplier methods it develops similarly that λ has a very basic economic significance.

Returning to the original policy of the company we have another way of indicating its irrationality. In every one of our optimal policies we can take any one of the items individually and it will have the given ratio between its orders per year and its average inventory. As a matter of fact, this is the essence of a rational inventory policy in a nutshell—making every item's policy consistent in this respect. Now, if we look at the current policy of the company we will find that there is no such uniformity among the individual items. This has nothing to do with the fact that the company policy isn't optimal for its own specific costs. It is that the policy could not be optimal for any conceivable single set of costs. For the company's current policy we can use our argument on the individual item policies and discover the imputed carrying costs assuming that the order cost is actually \$10:

ITEM	(T.I.)/(T.O.)	IMPUTED C_c
1	6.25	1.60
2	31.25	.32
3	41.67	.24
4	208.33	.05
5	62.50	.16

The wide range in these imputed values is an indication of the irrationality of the company's current policy.

Finally, our argument provides the means for a device which can be most useful to the executive responsible for establishing inventory policies under circumstances where it is difficult or impossible to obtain satisfactory estimates of the relevant costs. This device is called the optimal policy curve. The basis of this curve is precisely the constant product, $(T.I.)(T.O.)$, which we have found holds true, for a given set of data, for the optimal policy for any possible combination of values of C_r and C_c . We can make a graph of this curve, using total inventory investment as one axis and total orders per year for the other axis. The equation, $(T.I.)(T.O.) = K$, is that of a hyperbola. Figure 4-2 shows this hyperbola, the optimal policy curve, for the data of our example. Every point on this curve represents the optimal policy for some ratio of C_r to C_c , specifically, the ratio for any point on the curve is given by $(T.I.)/(T.O.)$. Further, every possible optimal policy is represented by a point on this curve.

By means of this curve we can recapitulate our argument. Point *A* is the current company policy. Its irrationality is indicated by the fact that it does not fall on the curve. When we argued that we could improve the company position by holding orders at 60 and reducing the inventory investment, we were simply saying that by running a parallel to the T.I. axis from point *A* until it intersects the optimal policy curve, at *B*, we must be improving the

situation. When we held the inventory fixed at \$4200, we ran a parallel to the T.O. axis from *A* until it intersected the optimal policy curve at *C*. This, obviously, must also improve the company position.

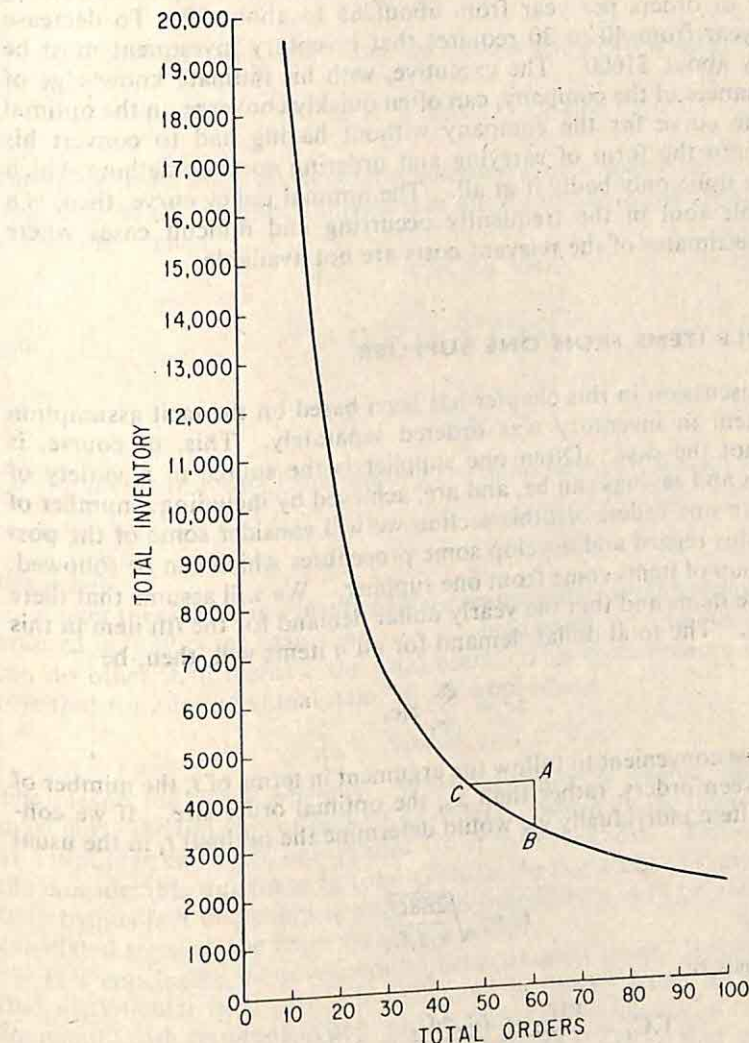


FIG 4-2 Optimal policy curve.

However, the optimal policy curve can be considerably more useful than simply as a geometrical analogy to an already completed mathematical argument. It provides means for the executive to utilize his knowledge and experience under circumstances where this knowledge and experience cannot be boiled down into the form of cost estimates. Thus, the optimal policy

curve shows the executive exactly how orders and inventory investment can be traded one for the other. For example, it can be seen immediately that to reduce the inventory investment from \$3000 to \$2000 will require increasing the number of orders per year from about 65 to about 99. To decrease orders per year from 40 to 30 requires that inventory investment must be increased by about \$1600. The executive, with his intimate knowledge of the circumstances of the company, can often quickly converge on the optimal point on the curve for the company without having had to convert his knowledge into the form of carrying and ordering costs—something which can often be done only badly if at all. The optimal policy curve, then, is a most valuable tool in the frequently occurring and difficult cases where satisfactory estimates of the relevant costs are not available.

32. MULTIPLE ITEMS FROM ONE SUPPLIER

All of our discussion in this chapter has been based on the tacit assumption that each item in inventory was ordered separately. This, of course, is frequently not the case. Often one supplier is the source of a variety of related items and savings can be, and are, achieved by including a number of such items on one order. In this section we will consider some of the possibilities in this regard and develop some procedures which can be followed.

Let a group of items come from one supplier. We will assume that there are n of these items and that the yearly dollar demand for the i th item in this group is $z_i c_i$. The total dollar demand for all n items will, then, be

$$zc = \sum_{i=1}^n z_i c_i$$

It will be most convenient to follow the argument in terms of t , the number of months between orders, rather than x_0 , the optimal order size. If we consider the i th item individually we would determine the optimal t_i in the usual way and get

$$t_i = \sqrt{\frac{288C_r}{c_i z_i C_c}}$$

with total costs of

$$T.C_i = \frac{12C_r}{t_i} + \frac{t_i z_i c_i C_c}{24} = \sqrt{2C_r C_c z_i c_i}$$

Alternatively, we could consider all the items in the group as representing one unit—stored, so to speak, on different shelves—and calculate the ordering policy for this one unit. We would find

$$t = \sqrt{\frac{288C_r}{C_c \sum_{i=1}^n z_i c_i}} = \sqrt{\frac{288C_r}{C_c zc}}$$

with total costs given by

$$T.C. = \sqrt{2C_r C_c} \sqrt{\sum_{i=1}^n z_i c_i} = \sqrt{2C_r C_c} \sqrt{zc}$$

The total costs for the n items when each one is optimized separately would be

$$\sum_{i=1}^n T.C_i = \sqrt{2C_r C_c} \sum_{i=1}^n \sqrt{z_i c_i}$$

The total costs when all items are considered as a unit is obviously much smaller. As an example, suppose that all of the $c_i z_i$ in the group are equal to some K . Then

$$T.C. = \sqrt{2C_r C_c} \sqrt{nk}$$

and

$$\sum_{i=1}^n T.C_i = (\sqrt{2C_r C_c})(n)\sqrt{k}$$

so

$$T.C. = \frac{1}{\sqrt{n}} \sum_{i=1}^n T.C_i$$

The total cost resulting from treating the items as a unit would, in this case, be only $\frac{1}{\sqrt{n}}$ times the sum of the items' individual total costs when treated separately.

Under these assumptions there is no reason whatsoever for any item to be ordered on any other basis than $t_i = t$. No change of t_i to any other value can do other than increase the total costs. This is so because it is always true that for any individual item, A , we would have

$$\sqrt{z_A c_A} + \sqrt{zc - z_A c_A} > \sqrt{zc}$$

The reason for this is obvious. When each item is treated individually it must carry the full cost of each of its orders. When all the items are treated as a unit then each item only carries $1/n$ of the order cost. This accounts for the considerable difference in total costs in the two cases and also results in the obvious fact that every t_i , calculated individually, will be greater than t , calculated from all the items treated as a unit.

This conclusion, however, contradicts common sense. It would appear that a particular item with a low demand in dollars could be ordered less frequently with resultant savings in total costs. Specifically, we could order such an item only at some integral multiples of t , say rt , and there should certainly be savings as compared to ordering it every t months, assuming that the total dollar demand for the item is small enough. However, the assumptions we have made permit no such saving for an individual item because there are no costs in the total cost equation which depend directly on the number of items on the order. It is this which is contrary to common sense. There is certainly a fixed cost for one order which is independent of the number of items on the order. The cost of preparing the order includes such components and, in addition, the cost of paying for the order is such a fixed cost.

However, it is equally true that there are some costs which depend on the number of items on the order. Specifically, each item, to be included on the order, must be reviewed—sometimes a physical inspection, more often pulling out a stock card and making some calculations. This review takes time and, therefore, has a cost and this cost varies with the number of items on the order. If we want to accomplish an analysis that will be in accord with common sense—and reality—we must introduce separate costs to represent these different components. For this purpose we will let C_{rF} represent the fixed cost of the order. For mathematical convenience we will express the review cost for the individual item as a multiple of this fixed cost, hC_{rF} . We will assume, as a reasonable approximation, that h is constant for all items in the group. The introduction of this latter cost does not change the form of the total cost equation:

$$T.C_i = \frac{t_i z_i c_i C_c}{24} + \frac{12(1+h)C_{rF}}{t_i}$$

which is minimized by

$$t_i = \sqrt{\frac{288(1+h)C_{rF}}{z_i c_i C_c}}$$

with minimum total costs for the item of

$$T.C_i = \sqrt{2C_c(1+h)C_{rF}z_i c_i}$$

and with total costs for all items, optimized individually, of

$$\sum_{i=1}^n T.C_i = \sqrt{2C_c(1+h)C_{rF}} \sum_{i=1}^n \sqrt{z_i c_i}$$

Now, if we consider all n items as a unit we have

$$T.C. = \frac{tzcC_c}{24} + \frac{12C_{rF}}{t} + \frac{12nhC_{rF}}{t}$$

This is minimized by

$$t = \sqrt{\frac{288C_{rF}(1+hn)}{zcC_c}}$$

with total costs of

$$T.C. = \sqrt{2C_c C_{rF}(1+hn)} \sqrt{\sum_{i=1}^n z_i c_i}$$

Comparing the two approaches we see that the situation is more complex than it was when we ignored the hC_{rF} component. For example, a specific t_i can now be smaller than t . This will occur if

$$\frac{1+h}{1+hn} < \frac{z_i c_i}{zc}$$

Frequently, h will be small. If, for example, $h = .05$ and $n = 20$ it would follow that if one item accounted for more than 52.5% of the total dollar

demand of the whole group of items then t_i for that item would be less than t .

The effect of the introduction of the hC_{rf} costs associated with the review of the individual items is to create two possibilities which did not arise when we did not include these costs. First, it is possible that using t for all items will not minimize total costs if one item has a large percentage of the total dollar demand for the group of items. Second, total costs may be lowered, even when t is optimal, by ordering items with small demand at integral multiples of t . We will consider these possibilities in order.

Suppose one item, A , accounts for a large proportion of the total dollar demand for the group of items, say $z_A c_A = azc$. Then the total dollar demand of the other items will be $(1 - a)zc$. Suppose, now, that we determine t_A for the single item separately:

$$t_A = \sqrt{\frac{288(1+h)C_{rf}}{azcC_c}} \quad \text{and} \quad T.C_A = \sqrt{2(1+h)azcC_cC_{rf}}$$

Designating all the remaining $(n - 1)$ items by B , we can determine t_B for this group taken as a single unit:

$$t_B = \sqrt{\frac{288[1 + (n - 1)h]C_{rf}}{(1 - a)zcC_c}}$$

and

$$T.C_B = \sqrt{2[1 + (n - 1)h](1 - a)zcC_cC_{rf}}$$

We want to find the conditions under which the total costs thus determined will be less than the total cost when t is found and used for all items:

$$T.C_A + T.C_B < T.C.$$

$$\text{or} \quad \sqrt{(1+h)a} + \sqrt{[1 + (n-1)h](1-a)} < \sqrt{1+nh}$$

$$\text{or} \quad a^2(2+nh)^2 - a(2nh^2 + 4nh + 4) + h^2 > 0$$

This quadratic inequation holds when

$$a > \frac{nh^2 + 2nh + 2 + 2\sqrt{(nh+1)^2 + h^2(nh+1)(n-1)}}{(nh+2)^2}$$

Assuming $h = .05$ we can construct the following table as an example of this equation:

n	a
3	.9978
4	.9954
5	.9921
10	.9676
20	.8997

The meaning of this table is the following: For $h = .05$ and $n = 20$ items in the group supplied by one supplier, unless one item represents more than 90% of the total dollar demand for the group of items it is correct to determine t for the entire group.

These calculations were based on the assumption that both A and B carried their full order costs. It would be expected that further savings could be achieved by setting the order period for B , t_B , at the nearest integral multiple of t_A and thus saving the $(12C_{rF})/t_B$ cost but not, of course, the $12(n-1)hC_{rF}/t_B$ cost. Further, there is no reason to confine the A "group" to one item. Let us, therefore, relax both of these assumptions. Let the items be divided into two groups, A and B , such that group A has total dollar demand of azc with g items and, hence, group B has total dollar demand of $(1-a)zc$ with $(n-g)$ items. Then:

$$t_A = \sqrt{\frac{288(1+gh)C_{rF}}{azcC_c}} \quad \text{and} \quad T.C_A = \sqrt{2(1+gh)azcC_cC_{rF}}$$

We intend to use some integral multiple of t_A for t_B but the algebra gets rather heavy, so instead of going through all the computations we will be satisfied with an approximation which is sufficiently good for practical applications. In determining the condition under which it will be advantageous to follow the course we are analyzing we will use the minimizing value of t_B , calculated without the $12C_{rF}/t_B$ cost. This gives:

$$t_B = \sqrt{\frac{288(n-g)hC_{rF}}{(1-a)zcC_c}} \quad \text{and} \quad T.C_B = \sqrt{2(1-a)(n-g)hzcC_cC_{rF}}$$

Our condition is that

$$T.C_A + T.C_B < T.C.$$

and this can be simplified to

$$\sqrt{a(1+gh)} + \sqrt{(1-a)(n-g)h} < \sqrt{1+nh}$$

This inequality is satisfied when the following simple condition obtains:

$$a > \frac{gh+1}{nh+1}$$

Whenever this is not true it is more economical to use t for all items. If this is true for some g then t_A should be determined for A , containing those g items, and the nearest integral multiple of t_A should be used instead of the t_B calculated in this argument. This, then, answers the first question raised above.

The second question concerned the possibility that a specific item might contribute such a small dollar demand to the group of items that it would be more economical to use some integral multiple of t as its order period rather than t itself. We will consider an item, D , which has relatively small dollar demand. We have determined t for all the items in the group and if t is used for item D then the total cost for this item will be

$$T.C_{D,t} = \frac{t z_D c_D C_c}{24} + \frac{12 h C_{rF}}{t}$$

The smallest possible change we could make in the order period for D would be to $2t$. In this event the total costs for D would be

$$T.C_{D,2t} = \frac{2tz_Dc_D C_c}{24} + \frac{12hC_{rF}}{2t}$$

The change to $2t$ would be advantageous only if

$$T.C_{D,t} - T.C_{D,2t} = -\frac{tz_Dc_D C_c}{24} + \frac{6hC_{rF}}{t} > 0$$

$$\text{or} \quad \frac{144hC_{rF}}{z_Dc_D C_c} > t^2 = \frac{288C_{rF}(1 + hn)}{zcC_c}$$

$$\text{or} \quad \frac{h}{1 + nh} > \frac{z_Dc_D}{zc}$$

Since z_Dc_D/zc is simply the percentage of the group dollar demand which is represented by the given item's dollar demand this is an easy criterion to apply. For example, with $h = .05$ and $n = 20$ this criterion says that we should use $2t$ as the order period for any item which has less than $1\frac{1}{4}\%$ of the total dollar demand of the group.

Of course, there is no reason why we should not use some integer larger than two as the multiplier of t . Let the integral multiple of the overall order period be jt . Then a repetition of our previous argument shows that the smallest j should be used for which

$$\frac{z_Dc_D}{zc} < \frac{h}{(j-1)(j)(1 + nh)}$$

For $h = .05$ and $n = 20$, we find from this expression that $3t$ would be used if $\frac{z_Dc_D}{zc} < .4\%$ and $4t$ would be used if $\frac{z_Dc_D}{zc} < .2\%$. These various equations

and expressions, then, provide the necessary means for determining the optimal policies when more than one item can be put on one order. We will note here for future reference* that the potential savings resulting from the incorporation of many items on one order are generally quite large. Of course, most companies which are able to amalgamate items on single orders automatically do so. However, they do not necessarily do so in the optimal way, so further savings can be accomplished by utilizing the methods we have been developing in this chapter. The point for future reference is that, since this is so, any theoretical inventory system which does not permit such multiple item orders is incurring a cost of the kind we called, in the first chapter, systemic costs.

* We will again consider the question of grouping items when we analyze the criteria for the design of an inventory study. The arguments presented here will also play a part in the resolution of problems concerned with data-processing. See pages 181-190 and pages 261-266.

33. MULTIPLE WAREHOUSES

Many organizations maintain more than one warehouse for storing their inventories. The reasons for this are many—and obvious, so we will not discuss them. From our more specialized point of view we want only to raise the question whether, when there is more than one warehouse, any new possibilities or problems arise in terms of optimal inventory policies. Certainly it is possible for each warehouse to determine its own optimal policies exactly as if the warehouse were an independent company. Further, this is the way many companies manage their multiple warehouses. When this is so there is nothing more to be said than has been said above. However, frequently it is the case that companies with multiple warehouses have central ordering. This means that each warehouse sends periodic summaries of its needs to a central ordering office which then processes the orders for all of the warehouses. By amalgamating several warehouses' needs for the same item it is often possible to accomplish various kinds of savings in the form of discounts or transportation savings. We do not propose to enter into a discussion of the relative merits of centralized and decentralized ordering. It is sufficient that centralized ordering with multiple warehouses is the system which is used by a large number of companies. Our question is: How does this affect the optimal inventory policies?

Fortunately, we do not need to develop a new argument in order to answer this question. We have already answered it! A moment's thought suffices to show that the problem is exactly the same in structure as was the problem of multiple items on one order. Here the question is one of multiple warehouses on one order. The same specific questions arise as we treated in the preceding section and the answers are, of course, identical. C_{rF} for centralized ordering is the cost of processing one order in the central office. In addition, each individual warehouse has a cost of sending its requisitions to the central office. This is the cost we expressed as hC_{rF} . If it can be assumed that the cost of the item is the same for each warehouse there would be some simplification in the required analysis but, in any event, the conclusions would remain the same.

We will therefore leave it to the reader to develop the argument for warehouses if he so desires. Here we will only note that the savings that result from the amalgamation of several warehouses' needs on one order are of the nature of systemic costs. The design of the inventory system determines the extent to which systemic costs will be minimized.

PROBLEMS

1. (a) For a given item there is a constant demand rate. Yearly demand is 6000 units. The price of the item is \$5 per unit. The ordering cost is \$15 and the carrying cost is 12% per year. What is the optimal ordering policy?
- (b) To what order period should this be rounded off?
- (c) The manufacturer of the item offers a 1% discount if 1200 units are ordered. Should this discount be taken?

2. For a particular item there is a known seasonal pattern of demand:

MONTH	DEMAND
1	50
2	70
3	120
4	80
5	60
6	40
7	30
8	50
9	90
10	140
11	80
12	60

The order cost is \$10 and the carrying cost is 1% per month. The price of the item is \$6. Assuming that this demand pattern will be repeated, year after year, what is the optimal ordering policy?

3. A company maintains inventory on three items with the following characteristics:

ITEM	YEARLY DEMAND	PRICE
1	3600	\$20
2	24000	5
3	600	25

Current policy is to order each item once each month.

- What is the optimal ordering policy which gives the same total number of orders?
- What is the optimal ordering policy which gives the same total average inventory?
- What is the equation of the optimal policy curve?
- If order cost is \$10 and the carrying cost is 12% per year, what is the optimal policy?
- Assuming the order cost is \$10, what is the imputed carrying cost for each item under the current policy?

dynamic inventory problems under risk

34. GENERAL CHARACTERISTICS

Dynamic inventory problems under risk are characterized by two features: there is the possibility of a number of orders and there is a known probability distribution of demand. This, at least, is a rough description of this kind of inventory problem. Actually, since time is implicit in these problems we must specify the probability distribution of demand in terms of some interval of time. However, we will discuss this question in greater detail subsequently and be satisfied here with the rougher statement. It is obvious that dynamic inventory problems under risk or under uncertainty include the great majority of practical inventory problems. All items which are regularly used over a reasonable length of time would fall into this category unless the demand is known with certainty, which is relatively infrequent. Even items with a short demand lifetime—a popular phonograph record, for example—will fall into this category if more than one order is possible. Most department store and supermarket inventories are of this general type—as are the inventory problems of the manufacturers who supply them. Examples need not be multiplied because a few moments' thought suffices to establish the fact that the inventory problems treated in the three preceding chapters are the exceptional ones. The ones covered in this chapter and the next are more nearly typical of the run-of-the-mill practical problem.

All of the major kinds of costs can occur in this type of inventory problem. The ordering cost is involved since alternative frequencies of ordering are possible. As the other half of this possibility of varying the order frequency we have the carrying costs. Since there is a probability distribution of demand there is the possibility of not having sufficient stock. Therefore, we

must have the understock cost involved in our equations. In some problems of this type—particularly, items with a highly seasonal demand—there will be the corresponding possibility of an excess stock with its corollary overstock cost. However, for many items there will not be any necessity to introduce an overstock cost because there is demand for essentially an unlimited period of time. It will be remembered that when we were using the understock cost we were balancing it against the overstock cost. Therefore, the question naturally arises: what balances the understock cost when there is no overstock cost? The answer to this question results immediately from a simple conceptualization of this kind of problem which may be helpful in providing insight into the mechanics of the analysis. We have, on the one hand, the ordering problem of the last chapter. This, so to speak, provides for the average demand level and requires the balancing of ordering costs against carrying costs for the average demand level. On the other hand we have the problem of maintaining an adequate reserve stock, over and above the stock for the average demand, to protect ourselves against fluctuations in demand. In determining the optimal amount of this reserve stock we take into account the cost of being out of stock, the understock cost. As counterbalance to this cost we have the fact that we must carry this reserve stock also. Therefore, the opposing cost to the understock cost, when there is no overstock cost, is the carrying cost of the reserve stock.

One major, new feature of this kind of inventory problem should be explicitly noted. It is that for the first time, the question of the time lag between order and delivery becomes important. In the previous models such a time lag might well have existed but either there was nothing which could be done about it or a simple and obvious compensatory adjustment sufficed to deal with it. In either event, there was no theoretical problem because of it. Such is not the case with dynamic inventory problems under risk. As one illustration of this fact it can be seen that there is no possibility for an understock—unless it is deliberately accepted—when there is no time lag between order and delivery of the item. Thus, for a low ordering cost there would be no need to carry any reserve stock at all because any shortages could be instantly rectified by another order. In this kind of inventory problem, then, the time lag must be taken explicitly into account.

A large proportion of the more complex mathematical analyses in the literature are devoted to one or another variant of this general kind of inventory problem. We cannot hope to present discussions of all of the models which have been developed since, even if space permitted, the level of mathematics required in many cases exceeds our self-imposed limitation for this book. We will be satisfied with the more modest goal of presenting the major approaches and suggesting some caveats concerning the applicability of some of the more advanced methods.

35. THE BASIC KINDS OF INVENTORY CONTROL SYSTEMS

Before we plunge into some of the necessary analysis it will be worthwhile to take an over-all view of the major kinds of inventory control systems which are,

or can be, used for dynamic inventory problems under risk. There are two major systems and the mathematical analysis needed in order to determine the optimal procedure—in terms of any one system—depends on the system desired. Naturally, we will be interested in comparing the two systems in terms of the over-all optimum but first we must characterize each system by itself.

The two systems can be most conveniently differentiated in terms of the variables under the control of the decision-maker faced with a dynamic inventory problem under risk. There are two such controllable variables: the frequency of ordering and the amount ordered. When we analyzed dynamic inventory problems under certainty we saw that these two variables are really only one. In other words, under certainty if the decision-maker decides to order a given quantity then the frequency of his orders is automatically determined, and vice versa. Now, this is still true of inventory problems under risk insofar as the average demand level is concerned but it is not true of the random fluctuations in demand. In other words, the decision-maker has his choice of how he absorbs the fluctuations in demand. He can do it either by varying the amount he orders or by varying the frequency with which he orders. These two possibilities result in the two systems in question.

The first system has a fixed order size and lets the frequency of ordering be determined by the fluctuations in demand. This system operates by determining some amount of stock, based on the time lag between order and delivery, which is the signal that another order of the given fixed amount must be placed. Whenever the stock on hand falls to this minimum level the given order is immediately placed. This system has been used in practice for a long time and is often called the two-bin system. The name stems from the fact that one convenient way to operate this system is to maintain two bins for each item. In the second bin is kept the minimum level and all the rest of the order is put in the first bin. Thus, when the first bin is emptied there is an automatic signal that the order must be placed and the required stock to last throughout the period between order and delivery is in the second bin. The major disadvantages of the fixed order quantity system is that it requires perpetual auditing of the inventory on hand. While this requirement may be partly alleviated for items which can be kept in two bins it is not so easy, or inexpensive, for many other kinds of items that cannot be conveniently stored in two bins. This system calls for a fixed order *quantity* so we will designate it as the *Q*-system.

The second system has a fixed order period and, therefore, lets the size of the order vary with fluctuations in demand. The order period is determined by analysis and then the amount of the item in inventory is reviewed at intervals equal to the order period. The amount that is ordered will vary in accordance with whether the amount in inventory is lower or higher than average. The average amount that will be ordered is, of course, determined from the analysis, and the fluctuations in demand around the average are added to this amount to determine the total size of the order. This system has the advantage that it only requires periodic review of the inventory situa-

tion, which makes it particularly convenient when central stock records are maintained. Since this system has a fixed order *period* we will designate it as the *P*-system.

These are the two major systems which are used in practice. They are typically of maximum administrative simplicity and this is a major point in their favor, particularly because of the problems attendant on the systemic costs we briefly discussed in the first chapter. However, it is by no means the case that either one of these systems need be the true optimal system for the inventory problem of a specific item. We must emphasize this point by a little discussion of it.

For a dynamic inventory problem under risk it is easy to ask, in words, what is the optimal inventory policy? However, it is unfortunately the case that the mathematical translation of this question into the form of an equation to be solved results in a very difficult kind of functional equation. Very little can be said generally about the solutions of this equation. In a famous paper, Dvoretzky, Kiefer, and Wolfowitz* proved that this equation has a solution—in other words, an optimal policy exists—but there are no practical procedures for finding the solution in a specific case. This immediately suggests the need for a different approach. Instead of demanding the optimal inventory policy, we might narrow the number of possibilities by considering only policies of particular—and generally simple—types. This was the approach used, in another famous paper, by Arrow, Harris, and Marschak.† They considered a particular set or subgroup of policies of the two-bin type—which we have called *Q*-systems—and they referred to them as (*s*, *S*) policies. Now, there are two main kinds of questions which can be asked about such a subgroup of all possible inventory policies. First, we can ask how the optimal policy of the subgroup can be determined. Second, we can ask under what conditions would the optimal policy for the subgroup also be the true optimal policy over all possible policies—not just over the policies of the subgroup. In other words, we would like to know the limitations on the form of the demand distribution and the various relevant costs which would ensure that the optimum policy of the given subgroup would be the true optimum over all possible policies. For those cases where it is relatively easy to determine the optimum policy for the subgroup and where we can simultaneously prove that the optimum policy for the subgroup is also the optimum policy among all possible policies, it is clear that we have succeeded in solving the general functional equation by taking something of a detour. This leads to considerable interest in the second question for all subgroups for which the first question has a reasonable answer.

When we add to the natural interest in the second question mentioned in the preceding paragraph the fact that the functional equation in question has become the happy hunting ground for many of the more mathematically oriented operations research theorists—and economists—it becomes clear why

* Aryeh Dvoretzky, J. Kiefer, and Jacob Wolfowitz, "The Inventory Problem," *Econometrica*, 20 (1952), pp. 187–222, 450–66.

† Kenneth J. Arrow, Theodore E. Harris, and Jacob Marschak, "Optimal Inventory Policy," *Econometrica*, 19 (1951), pp. 250–72.

the literature is replete with examples of proofs of optimality in the sense of the second question. Now, while such proofs are interesting they are only of the most minor consequence in practice. First, they are useless unless an alternative is available in case the subgroup optimum policy is not the true optimum over all policies. For example, if the only subgroup optimum that can be calculated happened to be of the (s, S) type it then would be of little consequence if someone demonstrated that for a particular item this policy was not the true optimum. Second, the proofs of optimality are generally developed under restrictive assumptions in one crucial regard—they assume no time lag between order and delivery. This is so rarely the case in practice that these proofs have little practical content. Third, and also crucial, all of these proofs of optimality ignore completely the kinds of costs we have called systemic and these are generally of overwhelming importance in choosing between inventory control systems. For these reasons we are not going to concern ourselves here with answers to the second kind of question.

A word about the name " (s, S) policies" is in order at this point. We are not going to use this name at all in our discussion. Our reason is that no other term in the literature seems to be so subject to contradictory usages as this one. Inspection of the literature shows the most various definitions and usages of this idea of " (s, S) policies." The present authors do not know who has misunderstood what but that somebody has misunderstood something follows directly from the contradictory statements that can easily be culled from the literature. Our own understanding—and perhaps we are to be numbered among those who should use the prefix "mis"—is that the (s, S) system is a straightforward two-bin system. As such it is completely characterized by two parameters: s and S . Knowledge of these, and nothing else, is sufficient to define the (s, S) system for a given item. It is, in short, precisely what we are calling the Q -system. Now, Arrow et al., in the article cited previously, modified this system in the sense that they considered cases where review of the inventory took place at regular intervals but orders only occurred if the stock were less than or equal to s on the occasion of one of these reviews. Now, this alteration results in the necessity of having three parameters to describe the system: the review period, s , and S . This makes the modified (s, S) policy a variant of our P -system and not vice versa, as is occasionally implied in the literature. The P -system is not a variant of the (s, S) system, by any meaningful definition of the latter. Furthermore, to call the three-parameter variant an (s, S) system seems to us to introduce additional ambiguity by running counter to the usual procedure in classificatory systems. In any event, whatever may be the clearest statement regarding the (s, S) system, we will avoid further confusion by ignoring it. The Q -system and the P -system, as we have defined them, include all the major usable inventory systems.

36. DEMAND PROBABILITY DISTRIBUTIONS

There are some major differences between what we must know about the probability distribution of demand in the dynamic case and what we had to

know about it in the static case. These differences deserve special comment. The same problem concerning the demand distributions exists in both of our systems so we must clarify it before proceeding to the discussions of the individual systems.

When we considered the static inventory problem under risk we needed only knowledge of one demand probability distribution. Specifically, we needed the probability distribution of demand for the entire period to be covered by the inventory. This is not so in the case of dynamic problems. Here we need to know what the probability distribution of demand will be for varying length time intervals. It is not sufficient, for example, to know the probability distribution of demand for a one-week period because we may, in our analysis, need to know it for three-week periods or for five-month periods. Now, it is obvious that there must exist some relationship between the probability distributions of demand for each week and those, for example, for each month, which are simply the sums of four weekly demands. Our problem is to describe this relationship mathematically.

In order to avoid analytical complexities it is customary in the literature to make a major assumption about the probability distributions of demand for consecutive time periods. Whatever the basic period chosen, say a week, for expressing the demand distribution it is assumed that the probability distribution for each period is the same. This means, in particular, that we are assuming that the demand is stationary and that demand in following periods is independent of the demand level in preceding periods. This assumption is a rather ruthless one but we will make it for pedagogical reasons. It should not be concluded that it is impossible to analyze inventory problems for which this assumption cannot be made but it is certainly much more difficult.

Granted this assumption, the question we are raising is: Given the probability distribution of demand for a time period of specific length, how can we determine the probability distribution of demand for some longer time interval composed of some integral number of these basic periods? Suppose, for example, that we have the following simple demand distribution for some item on a weekly basis:

DEMAND	PROBABILITY
0	.1
1	.4
2	.3
3	.2

What would be the probability distribution of demand for a two-week period? For such a simple distribution it is easy to calculate the new distribution. Clearly, for two weeks the demand cannot be less than zero. This will only happen if the demand in each of the two weeks is zero. This has a probability of $.1 \times .1 = .01$. The demand for two weeks cannot exceed six units and this will only happen if the demand in each of the two weeks is three units, which has a probability of $.2 \times .2 = .04$. To calculate

the probabilities of the intermediate levels of demand we proceed as follows, taking a two-week demand of three units as an example. There are four ways in which we can get a total demand of three units in two weeks:

DEMAND IN FIRST WEEK	DEMAND IN SECOND WEEK	PROBABILITY
0	3	$.1 \times .2 = .02$
1	2	$.4 \times .3 = .12$
2	1	$.3 \times .4 = .12$
3	0	$.2 \times .1 = .02$
		.28

The total probability of a demand for three units in two weeks is therefore .28. We could proceed similarly and painstakingly determine the probabilities of all possible levels of demand for two weeks and, indeed, we could extend the procedure directly to determining the probability distributions of demand for an arbitrary number of weeks—admittedly with considerable labor. Now, while we cannot avoid most of the computational work involved we can routinize it a great deal more so that less thinking is required. Rather than try to argue our way through the next step we will simply credit it to a flash of intuition on somebody's part. This intuitive flash is that we can accomplish the same thing as above by using binomial expansions. Thus, let us write our probability distribution for one week in the following form:

$$.1 + .4x + .3x^2 + .2x^3$$

In this expression the exponent of x equals the demand (for the first term $x^0 = 1$) and the coefficient is the probability of the level of demand represented by the exponent. The advantage of writing the demand distribution in this way is that if we want to know the probability distribution of demand for n weeks we need only raise this expression to the n th power and read the result using the same convention. Thus, the probability distribution of demand for two weeks is given by

$$(.1 + .4x + .3x^2 + .2x^3)^2 \\ = .01 + .08x + .22x^2 + .28x^3 + .25x^4 + .12x^5 + .04x^6$$

Our previously calculated probabilities check with the corresponding coefficients and all the other probabilities for two-week demands are given as well. Since it is not easy to raise such an expression to a higher power it is not easy to determine the probability distributions of demand for large values of n but it is at least a routinized procedure.

This process produces what are called, technically, the convolutions of the original probability distribution with itself. We used a discrete probability distribution as our example and it is obvious that we can, in theory at least, use exactly the same procedure for any discrete probability distribution. The same thing can also be done for continuous probability distributions. The method we used for discrete distributions of writing it in the form of a binomial expansion is technically known as the method of generating

functions. In the case of continuous distributions the analogical method is that known as characteristic functions. In either event, the result is to obtain the convolutions of the original distribution with itself. This result, obviously, is exactly what we need in order to go from information about the probability distribution of demand for a given fixed time period to the demand distribution for any integral multiple of that time period.

Unfortunately, the convolutions of an arbitrary probability distribution with itself, though perfectly well defined and calculable, are not generally very tractable in terms of the methods of mathematical analysis. This leads to a strong desire on the part of the analyst to select specific probability distributions which are extremely well-behaved in terms of their convolutions. Fortunately, there are a great many such distributions and numbered among them are the best-known probability distributions: the binomial, the normal, and the Poisson. It is this fact which will lead us to use frequently the normal distribution as an example in this chapter. Once again, however, it is desirable to emphasize that this does not mean that our ability to resolve dynamic inventory problems under risk stands or falls with the assumption of one of the better-known probability distributions as representing demand. Without this assumption it is simply harder, not impossible.

When we say that the convolutions of the mentioned distributions are particularly well-behaved we mean specifically that for each of these distributions its convolutions are the same kind of distribution. Thus, a binomial distribution convoluted with itself any number of times always gives a binomial distribution and similarly for the normal distribution and the Poisson distribution. If we add two important, and obvious, facts we can be more specific. The first fact is that the n th convolution of any stationary* distribution has a mean equal to n times the mean of the distribution. This is a commonplace. If average demand for one week is 10 units then, under our assumptions, the average demand for, say, five weeks will be simply $5 \times 10 = 50$ units. The second fact is not so intuitively obvious but it can easily be proved—although we will forgo proving it here. This is that the variance of the n th convolution of any distribution which meets our basic assumption will be equal to n times the variance of the distribution. Translating these two facts into terms of the normal distribution we end up with: the n th convolution of a normal distribution with mean of $\bar{z}$ and standard deviation of s will have mean equal to $n\bar{z}$ and standard deviation equal to $s\sqrt{n}$. The square root of n appears here, of course, because we stated the relationship in terms of the standard deviation rather than in terms of the variance. Thus, with no difficulty whatsoever we can determine the demand for any integral multiple of our basic time period if the demand distribution for the basic period is normal.

Finally, we need some notation to represent convolutions. We will use, in this and the following chapters, $f(y) dy$ to represent the probability distribution of demand in the basic time period. We will let $f^i(y) dy$ represent the probability distribution of demand convoluted i times. This is, then, the

* For a discussion with respect to the concept of stationarity, see pages 318–319.

probability distribution of demand for i times the basic time period. If $f(y) dy$ represents the probability distribution of demand for one week, then $f^i(y) dy$ represents the probability distribution of demand for i weeks.

37. THE Q-SYSTEM OF INVENTORY CONTROL

As we have already indicated, the Q -system has a fixed order size and a varying order period. The procedure is that whenever the stock on hand falls to a certain minimal level, based on the time lag between order and delivery of the item, an order is automatically placed for the predetermined fixed amount. Fluctuations in demand, then, are reflected by the resulting variations in the time between orders. We must now undertake to discover how an optimal system of this type can be determined for a given item. Such a system is completely determined by knowing the order size and the minimal stock level which represents the signal to place the order. It is these factors that we must find means to determine.

Let us restate the handling of demand fluctuations in the Q -system. Since these fluctuations are absorbed by varying the order period it follows that there is no need for reserve stock to meet these fluctuations. Leaving aside the period of time lag for a moment, we need only order the average amount required to meet the average demand for the order period. If fluctuations in demand result in this stock being exhausted more quickly, then the subsequent order will simply be placed more quickly. Therefore, there is no need for reserve stock to meet these fluctuations. This is not true of the lag time period since the order is already placed when the lag period begins and any demand in this period can only be met from stock on hand. Therefore, there must be a reserve stock to absorb any fluctuations in demand during the lag period. Finally, we note that even though the order period will vary in response to demand fluctuation it is still the case that the average order period is immediately determined by the fixed order size. This is evidently true since the average demand for any period, say a year, is known and if we divide this by the fixed order size we must get the average order period.

This suggests a very simple way for finding the desired policy. Let us use the optimal lot-size formula on average demand to determine the optimal number of orders per year and, hence, the fixed order size. Then we can separately determine the optimal reserve stock for the lag period by the methods of static inventory problems under risk. How would this work? As an example we will take an item with weekly demand normally distributed with mean of $\bar{z} = 50$ and standard deviation of $s = 5$. We will assume that the lag time (also called the lead time) between order and delivery is three weeks. For our cost data we will assume that the order cost, C_r , is \$10 and the carrying cost, C_c , is 12% per year and that the item price per unit is \$5. This is sufficient for the first part of our calculation. Yearly demand for this item is simply $52(50) = 2600$ units. Using the optimal lot-size equation

developed in Chapter 4 we can immediately determine the order size:

$$x_o = \sqrt{\frac{2(10)(2600)}{5(.12)}} = 294.4$$

This, then, would be the fixed order size and we can immediately determine the average number of orders per year as equalling $2600/294.4 = 8.83$.

For the three-week lead time period between order and delivery we need 150 units to meet average demand. This, however, is already built into the system in the sense that we shall simply order three weeks in advance to allow for this lag. Not yet built into the system is the need for a reserve stock to meet fluctuations in demand during the lead time period. The probability distribution of demand during a three-week period is given in accordance with our discussion of the preceding section. Since demand is normally distributed during one week we know that the distribution of demand will be normal for three weeks. Specifically, for our example the three-week distribution will be normal with mean of 150 and standard deviation of $s' = 5\sqrt{3} = 8.66$. As we know from the discussion of Chapter 4, we can determine the reserve stock either by knowing C_u or else by having management set a permissible percentage level of understocks—which, of course, imputes a value to C_u . For our present example let us suppose that management has set a performance standard requiring that there should be only one chance in a thousand of being understocked. From the tables of the normal distribution we immediately discover that this standard requires a safety allowance, or reserve stock of $3.085 s'$. In our case this means that we need $3.08(8.66) = 26.7$ units in reserve during the lag period. In other words, the total requirement for the lead time period is 150 units for average demand and 27 units for reserve or a total of 177 units. This, then, is our minimal amount of stock which serves as a signal for an order. Whenever the stock drops to 177 units we immediately order 294 units. This is our Q -system for the item under consideration.

Now let us calculate the total cost of the system we have developed. There are an average of 8.83 orders per year at a cost of \$88.30. The 27 units for reserve will sometimes be partly, or entirely, used up in the lead time period but equally often the lead time will end with more than 27 units in stock. Therefore, on the average the reserve stock will be carried all year round. This will cost $27(5)(.12) = \$16.20$. The carrying cost for the average demand will be simply one-half of the order size perpetually or a cost of $147(5)(.12) = \$88.20$. Thus, the total cost will be \$192.70. Of course, since we did not assume knowledge of C_u we cannot calculate the out-of-stock cost. This is presumably handled satisfactorily by the performance standard set by management.

Is this the minimum possible total cost for a Q -system for this item? Not at all! It will probably be reasonably close and therefore can be—and, in practice, is—used as a reasonable approximation to the true optimum. The reason that it is not optimal is that we have tacitly assumed something which isn't true. We have blithely calculated the optimal order period and then the

optimal reserve stock for the lead time period as if there were no interaction between them. The truth is, of course, that there is interaction between them and this interaction must be taken into account if it is desired to find the true optimal Q -system. The interaction results from the fact that the protection level cannot be meaningfully expressed in absolute terms as we expressed it in our calculation above. We stated that management demanded that there should only be one chance in a thousand of going out-of-stock. Now, if we had an order period of one year this would mean that, on the average, there would be one out-of-stock in one thousand years. If, instead, we had an order period of one week there would be one out-of-stock approximately every twenty years. Obviously, the economic implications of these two extremes are totally different and this points up the fact that the performance standard cannot be given in these absolute terms. Instead, it must be given on a time basis. Generally, for such a performance standard it will result that the longer the order period, the smaller, relatively, need be the reserve stock for the lead time period. It is this interaction which must be taken into account if the optimal Q -system for an item is to be determined.

In order to illustrate how we can proceed to determine the optimal Q -system, let us switch our example to the extent that we assume a knowledge of the out-of-stock cost. Let us assume that this is a fixed cost, independent of the number of units out of stock. It can easily be calculated that the performance standard set by management in our example imputes a fixed out-of-stock cost of \$168.10* and we will use this as the fixed out-of-stock cost in our continuation so that the two results will be comparable. With this imputed out-of-stock cost in our original example we can immediately calculate that the yearly out-of-stock cost would be $8.83(168.10)(.001) = \$1.48$. Thus, the actual total cost, as originally calculated, would be \$194.18 per year.

We can set up the necessary total cost equation to determine the optimal Q -system quite directly by simply enumerating the various cost components and summing them to form the total cost equation. If we have an average order period of t weeks the total costs per year will be:

$$\text{Ordering cost} = \frac{52C_r}{t}$$

$$\text{Carrying cost for average demand} = \frac{tDcC_c}{104}$$

where D = demand per year.

* This assumes an understock once in a thousand lead times. The ordinate of the lead time demand distribution is .0004042 at the given reserve stock. The imputed value of the fixed out-of-stock cost is given by

$$\frac{.60}{8.83(.0004042)} = \$168.10$$

Letting w represent the lead time period safety reserve and R the average demand in the lead time period:

$$\text{Carrying cost for safety reserve} = wcC_c$$

$$\text{Out-of-stock cost (per year)} = \frac{52K}{t} \int_{R+w}^{\infty} f(y) dy$$

where K is the fixed out-of-stock cost. Therefore,

$$\text{T.C.} = \frac{52C_r}{t} + \frac{tDcC_c}{104} + wcC_c + \frac{52K}{t} \int_{R+w}^{\infty} f(y) dy$$

This must be minimized simultaneously with regard to t and w :

$$\frac{\partial \text{T.C.}}{\partial t} = \frac{-52C_r}{t^2} + \frac{DcC_c}{104} - \frac{52K[1 - F(R+w)]}{t^2} = 0$$

$$\frac{\partial \text{T.C.}}{\partial w} = cC_c - \frac{52K}{t} f(R+w) = 0$$

These two simultaneous equations cannot be solved explicitly but they can be solved quickly enough by iteration. Solving each equation for t and equating the results we get

$$[f(R+w)]^2 = \frac{2cC_c\{C_r + K[1 - F(R+w)]\}}{DK^2}$$

In our example, $f(R+w)$ is the normal distribution of demand for three weeks, with mean = 150 and standard deviation = 8.66. In using the normal tables to solve this equation it must be remembered that the ordinate, $f(R+w)$, of a given distribution is equal to $\frac{1}{s}$ times the corresponding ordinate as found in the table. Thus, in our case—letting $g(w)$ represent the ordinate looked up in the table, we want

$$[g(w)]^2 = (8.66)^2 [f(R+w)]^2$$

or, using our specific values

$$[g(w)]^2 = \frac{10 + 168.1[1 - F(R+w)]}{816400}$$

This can be solved very quickly by assuming at the first step that $F(R+w) = 1$ and solving for $g(w)$. Then use this value of w to find the value of the right-hand side of the equation and solve for $g(w)$ again. Generally, no more than three such iterations will be required. We find

$$w = 3.08$$

This gives $g(w) = .0035$ and, hence,

$$f(R+w) = \frac{.0035}{8.66} = .000404$$

Therefore, from the $\frac{\partial T.C}{\partial w}$, equation, we have

$$t = \frac{52(168.10)}{.60} (.000404) = 5.89$$

Translated, our order size will be $5.89(50) = 294.5$ units. Our reserve stock will be $3.08(8.66) = 26.7$ units. Our procedure is to order 294 units whenever the stock level falls to 177 units.

It will be seen that we have rediscovered exactly the same Q -system as we determined from our approximate analysis. It really isn't "exactly" the same but it is within the usual rounding errors. The optimal policy will not always be this close to the approximate solution but it will generally be not very far from it. This is why the approximate solution is sufficient for most practical purposes. However, the exact solution can be obtained by the method we have developed and in any specific case the divergence of the approximate solution from the true optimum can be determined.

We could repeat the same procedure for an understock cost per unit, C_u , and similarly establish the optimal Q -system under these circumstances. The only change is in the one term in the total cost equation which expresses the out-of-stock cost. The procedure is the same but the derivatives of the total cost equation are more messy and there is more labor in determining the solution. Nonetheless, the extension to this case is straightforward. We will not give the equations here.

This, then, is the general procedure for determining an optimal Q -system. As will be seen in the next section, where we discuss the P -system, the optimal Q -system is generally less expensive than the optimal P -system for the same item. More specifically, it is true if we consider only the costs which enter into the respective total cost equations for the two systems. Consequently, it is particularly important that we emphasize the fact that there are other costs, not included in the equations, which can greatly change the picture resulting from a comparison of any of these different systems. In the first chapter we referred to these costs as systemic costs. We will not attempt to list all of the relevant systemic costs involved in the utilization of the Q -system but we can mention some of them. First is the already mentioned fact that a Q -system requires perpetual inventory review. This is obviously necessary since we must know as soon as an item's inventory reaches its minimal level so that the order for the item can be placed. Second, under the Q -system an item's order can occur at any time. This prevents the economies which result from amalgamation of several items from one supplier on one order. Third, and for the same reason, it is impossible under the Q -system to gain the economies of amalgamation of the orders of several warehouses into one central warehouse. Both of these potential savings were discussed in the last

chapter and it is clear that these savings can be of considerable magnitude. One must charge a cost to the *Q*-system since it makes it impossible to gain these savings. Our subsequent discussions in Part 2 consider these issues again in the light of systems analysis.

38. THE *P*-SYSTEM OF INVENTORY CONTROL

Under the *P*-system of inventory control there is a fixed order period and a varying order size. The procedure is that at periodic intervals—the period being analytically determined—the amount of inventory is reviewed and an order is placed. The total amount which should be on hand and on order is determined from the analysis so the amount of the order is determined directly by subtracting the amount on hand from this predetermined total. Incidentally, when the order period turns out to be shorter than the lead time period the inventory must be counted as including units on hand plus units on order but not yet delivered. This same procedure must be followed, obviously, in the case of the *Q*-system. The *P*-system is completely determined when it is known what the order period is and what the amount on hand and on order should be. These, then, are the two parameters which we must discover how to determine analytically.

The major difference between the *P*-system and the *Q*-system is that for the former all fluctuations in demand must be allowed for in determining the reserve stock. Since there is no flexibility in order period any fluctuation in demand must be met from stock carried for that purpose. In the *Q*-system we needed reserve stock only for the lead time period. In the *P*-system the situation is more complex. First, it is obvious that for the *P*-system it is necessary to maintain reserve stock as protection against demand fluctuations during the order period. The *Q*-system absorbed these fluctuations by variations in the order period and, hence, did not require reserve stock for this purpose. Second, however, is the fact that it will not be sufficient to have reserve stock in the *P*-system based only on the order period. To see why this is so it will be convenient to consider an example. Suppose that for a given item the order period and the lead time period are of the same length. Further suppose that we have, by the methods developed below, determined the reserve stock for the order period demand fluctuations. Now, at each order point the previous order will arrive in stock simultaneously with the determination of the next order. We will determine the amount to be ordered by comparing the existing stock level at the time of order with some determined amount which we want to be on hand at the beginning of an order period. Suppose, for example, that average weekly demand is 100 units, that the order period and lead time are each four weeks, and that the reserve stock, based only on the order period, has been found to be 20 units. Then we would want 420 units on hand at the start of each period. As long as demand remained exactly at its average level we would order 400 units every four weeks and we would always have 20 units in reserve stock. But suppose demand suddenly exceeds 420 units in one order period. At the

beginning of the subsequent period 400 units arrive from the preceding order but there is no reserve stock. Therefore, there is a 50% chance of also going out of stock in this subsequent period simply because we went out of stock in the preceding one. In other words, the inventory situation cannot be rectified until one intervening order period has elapsed. This enhanced probability of again going out of stock in the subsequent period must certainly be taken into account in our mathematical formulation. Of course, this same phenomenon holds true for order periods and lead times of any length. Setting the two equal, as in our example, was simply to make the situation easier to visualize. Further, the phenomenon in question does not result only when there is an out of stock. The point is that there is an interaction between our order for any given period and all subsequent periods. An exact solution for a *P*-system would require taking these interactions explicitly into account. The ideas involved in doing this are presented in Section 40 in a somewhat different context. Because of the mathematical complexities involved in considering these interactions it is customary to make a simplification in practice which we will follow in *all* cases where we deal with *P*-systems. This simplification consists in assuming that a satisfactory system will be achieved if we calculate the reserve stock for the order period plus the lead time rather than for the order period alone.

As a first approximation to the optimal *P*-system we can proceed as we did for the *Q*-system—use the optimal lot-size formula to determine the optimal order period and then use essentially the methods of static inventory analysis to determine the optimal amount to order. This will not give nearly as good a first approximation as it did for the *Q*-system because the interaction between the two variables is much greater here. However, it is easier to do and it generally represents a considerable improvement over the typical intuitive kind of inventory policy used by companies in lieu of analysis. Therefore, the approximation is frequently employed even though it is not optimal.

As an example of this approximation we will use the same item as in the preceding section. For that item we determined from the optimal lot-size formula that the order period should be 5.89 weeks (8.83 orders per year). We want to determine the reserve stock for the order period plus the lead time, which is three weeks. This means that we need to know the demand distribution for $5.89 + 3 = 8.89$ weeks. Now, we have seen in Section 36 that by utilizing convolutions we can determine the demand distribution for an integral multiple of a period for which the demand distribution is known. We could, therefore, round our order period off to 9 weeks and then write the demand distribution based on the ninth convolution of the weekly distribution. Instead, as a better approximation, we will simply assume that we can utilize the same procedure that we used for integer multiples for fractional multiples of the basic demand distribution. In this way we will compute the demand distribution for the order period we have determined. The reserve stock will be determined by balancing the carrying costs against the out-of-stock costs. As a reasonable approximation of the carrying costs we can assume that the reserve stock will be carried constantly. Designating the reserve stock for

the order period by r , we want to find that r which minimizes the total cost equation. Carrying costs will be, per year, rcC_c . Out-of-stock costs we will assume first to be constant, regardless of the number of units out-of-stock. Then we can write

$$T.C_r = rcC_c + nK \int_{\bar{z}+r}^{\infty} f^t(y) dy$$

where $f^t(y)$ = t th convolution of $f(y)$,

t = order period plus lead time expressed in multiples of the basic period for which $f(y)$ is known,

n = number of orders per year,

K = fixed understock cost,

$\bar{z}$ = average demand for order period.

To minimize this with regard to r we proceed as usual

$$\frac{dT.C_r}{dr} = cC_c - nKf^t(\bar{z} + r) = 0$$

or,
$$f^t(\bar{z} + r) = \frac{cC_c}{nK}$$

For our example $f^t(y)dy$ is simply the normal frequency function with mean = $8.89(50) = 444.5$ and standard deviation = $5\sqrt{8.89} = 14.91$. For our data we want

$$f^{8.89}(\bar{z} + r) = \frac{.60}{8.83(168.1)} = .0004042$$

Reference to the normal tables shows that this equation is solved for $r = 2.90$ (14.91) = 43.2 . This means that for this item and with this approximation the amount on hand after the arrival of an order should average $294.5 + 43.2 = 337.7$. Thus, every 5.89 weeks the inventory of this item should be reviewed. Each time we review the stock position there will be included in the stock on hand 150 units to meet average demand in the lead time period. This must have no effect on the fact that we want 331 units in stock upon the arrival of each order. Therefore, our rule is to order the difference between 481 units and the stock on hand at each review. The effect of this rule will be that we will order somewhat less if demand fluctuations have been in the direction of smaller demand and somewhat more in the other case. This, of course, is the essence of the P -system. If the total cost of this procedure is calculated it is found to be \$205.10, to be compared to the corresponding \$194.44 total yearly cost of the optimal Q -system. As was to be expected, the cost of the P -system is somewhat higher. Of course, this is not really a fair comparison because we have not yet determined the optimal P -system.

We can equally well introduce a per unit out-of-stock cost, C_u , into our approximate analysis. Using the same notation as before we have

$$T.C_r = rcC_c + nC_u \int_{\bar{z}+r}^{\infty} (y - \bar{z} - r) f^t(y) dy$$

and setting the derivative with respect to r equal to zero gives

$$F^t(\bar{z} + r) = 1 - \frac{cC_c}{nC_u}$$

This is a simple equation to solve. For example, suppose the per unit out-of-stock cost, C_u , is \$10 for the item we have been considering. Substituting our data into the above expression gives

$$F^t(\bar{z} + r) = .9932$$

Reference to the normal tables shows that we must have $r = 2.47(14.91) = 36.8$ and the rest of the reasoning will proceed as above.

In both of these approximations we have been calculating the carrying costs for the reserve stock under the assumption that the reserve stock was held in stock constantly. This is a sufficiently good approximation for low price items but may not be so good for expensive items. We can easily eliminate this assumption if it is so desired. In fact, we know that sometimes the reserve stock will be partially used up and when this happens the carrying costs should not be calculated to include the part of the reserve stock that was used. If we want to take this fact into account we need to calculate the total cost of the reserve stock in terms of the average amount of reserve stock which will be in stock. So doing, gives the total cost equation for a fixed out-of-stock cost:

$$T.C_r = cC_c \int_{\bar{z}}^{\bar{z}+r} (\bar{z} + r - y) f^t(y) dy + nK \int_{\bar{z}+r}^{\infty} f^t(y) dy$$

Taking the derivative with respect to r and setting it equal to zero gives:

$$f^t(\bar{z} + r) = \frac{cC_c}{nK} [F^t(\bar{z} + r) - F^t(\bar{z})]$$

For our example $F^t(\bar{z}) = \frac{1}{2}$. This equation must be solved iteratively but it is relatively easy to do so because $F^t(\bar{z} + r)$ will generally be nearly equal to one and this can be used for the first step in the iteration. Using our data, we want

$$f^t(\bar{z} + r) = .0004042[F^t(\bar{z} + r) - \frac{1}{2}]$$

Solving this equation shows that $r = 3.13(14.91) = 46.7$. As compared to the corresponding case where we assumed a fixed reserve stock carrying cost we see that the reserve stock has gone up. This is in accord with expectations since the effect of our change has been to decrease the expected carrying cost of the reserve stock and this naturally results in an increase in the reserve stock and a corresponding decrease in the expected out-of-stock cost.

We can, of course, include this change in the calculation of the reserve stock carrying cost along with the assumption of per unit out-of-stock costs. So doing gives the total cost equation

$$T.C.r = cC_c \int_{\bar{z}}^{\bar{z}+r} (\bar{z} + r - y)f'(y) dy + nC_u \int_{\bar{z}+r}^{\infty} (y - \bar{z} - r)f'(y) dy$$

Setting the derivative of this with respect to r equal to zero gives

$$F'(\bar{z} + r) = \frac{cC_c F'(\bar{z}) + nC_u}{cC_c + nC_u}$$

For our example this becomes

$$F'(\bar{z} + r) = .9966$$

The solution to this equation is $r = 2.71(14.91) = 40.4$. The rest of the reasoning would go exactly like our first example in this section. Once again it will be seen that the effect of the change in method of calculating the reserve stock carrying costs has been to increase the reserve stock.

So much for various ways of approximating the optimal P -system for an item. What can we do about finding the true optimum? This is a somewhat messy business and we will only illustrate the general procedure with one example. We will take the first of the four preceding approximations as our example. Thus, we assume that the reserve stock cost is fixed and that the out-of-stock cost is also fixed. What we want to do is to write a total cost equation which shows how costs vary with regard to both the order period and the reserve stock. Then we will minimize this equation with regard to both variables simultaneously. We will write the equation in terms of x , the amount ordered, and r , the amount of reserve stock. We will let LT represent the lead time. It is simple enough to write the total cost equation:

$$T.C. = \frac{52\bar{z}C_r}{x} + \frac{xcC_c}{2} + rcC_c + \frac{52\bar{z}K}{x} \int_{(\frac{\bar{z}}{x} + LT)\bar{z}+r}^x f^{\frac{\bar{z}}{x} + LT}(y) dy$$

In this equation $\bar{z}$ is the average weekly demand. We are assuming that the demand distribution is known for a basic one-week period. The four terms on the right-hand side of the equation are, respectively:

- (1) Yearly ordering costs.
- (2) Yearly carrying costs for the stock to meet average demand.
- (3) Yearly carrying cost of the reserve stock.
- (4) Yearly out-of-stock cost.

The function under the integral, $f^{\frac{\bar{z}}{x} + LT}(y) dy$, is the $(\frac{\bar{z}}{x} + LT)$ th convolution of the basic demand distribution. If x units are ordered the order period will be

* "True" only within the bounds of our assumption that it is sufficient to calculate the reserve stock for the order period plus lead time, as discussed earlier in this section.

$x/\bar{z}$, in weeks, so we need the demand distribution for this number of weeks plus the lead time. This is given by the indicated convolution of the basic distribution.

The difficulty is that not much can be done with this equation without some more knowledge about the form of the demand distribution's convolutions. In our example we have been assuming a basic normal demand distribution. Taking advantage of this assumption and going through some annoying calculations given in the Appendix, we find that the optimal P -system, under the stated conditions, has $x = 280.6$ and $r = 35.4$. This is to be compared with our first approximation, for which $x = 294.5$ and $r = 36.1$. The total cost of our optimal P -system will be \$200.26, as compared to the \$200.49 cost of our approximation and the \$194.44 cost of the optimal Q -system.

By similar methods the optimal P -system corresponding to any of our other three approximations can be determined. In each case the mathematics is complicated and an iterative procedure must be followed to obtain numerical solutions. The one requirement is that we must be able to write the convolutions of the basic demand distribution in explicit form. If we cannot do this we must use approximate methods* to achieve a near-optimal P -system.

These, then, are the methods by means of which the optimal P -system can be determined or approximated. It will be noted that the P -system does not suffer as badly from some of the systemic cost deficiencies mentioned with regard to the Q -system. It is a comparatively straightforward matter to amalgamate orders from various warehouses if this is so desired, and it is relatively easy to use the methods of the preceding chapter to amalgamate orders for various items from the same supplier. The P -system is well-suited for central inventory control in general. Some of the systemic costs associated with the P -system will be discussed in the second part of this book. For the moment it is sufficient to note that there are compensating advantages of the P -system to offset, at least in part, the otherwise clearcut superiority of the Q -system.

39. AN ALTERNATIVE FORMULATION OF INVENTORY POLICIES: SWITCHING MATRICES

In both of the two preceding sections we have derived optimal policies for subgroups of the total group of possible policies. The policies might or might not be optimal when all possibilities are considered. All we know is that they are optimal for the stated subgroup. In other words, the P -system we derive is the best P -system but it may not—and probably will not—be the best possible system. We already discussed in Section 35 the fact that the question of discovering the optimal policy among all possible policies is an exorbitantly difficult one. Nonetheless, it will be useful to consider an approach which does permit the discovery of the optimal policy in the widest sense. This will enable us to show that totally different policies than the

* The method employed in Part 2 is a particularly simple approximation. It requires management specification of allowable outages. See pages 236–237.

ones we have considered heretofore may be optimal and it will also provide a conceptualization of inventory policies which is of great intrinsic interest. This section will present the new way of thinking about an inventory policy and the next section will then develop the method for discovering the optimal policy, based on this section's structuring of the problem.

The approach we are going to use is based on the fact that when an inventory policy is specified the behavior of inventory becomes a function only of the probability distribution of demand and the given policy. The inventory level, as it moves through time, is a stochastic process and we propose to study this stochastic process per se. Under many circumstances the stochastic process is of a particular sort: a Markovian process. When this is the case we have powerful tools from probability theory to enable us to investigate the behavior of the inventory system. In the discussion of such a stochastic process we usually refer to states of the system and the transitions from one state to another. In the case of inventories the states of the system are ordinarily simply the stock levels. The stochastic process of the inventory will be Markovian if the following holds true: given the state of the system at time t then no additional information concerning the states of the system at prior times will change the probability that the system will be in any given state at some future time. In other words, all necessary information about the system's behavior is contained in the knowledge of its current state. How it got to that state has no effect on our ability to predict future states. The example we will use here is Markovian and we will not further discuss any questions concerning non-Markovian systems.

We can proceed most simply by taking an example. Suppose we have an item with the following probability distribution of demand on a weekly basis:

DEMAND	PROBABILITY
0	.1
1	.4
2	.3
3	.2

We select such a simple distribution to avoid being overwhelmed by arithmetic. We will further suppose that there is a one-week lag in delivery. Orders are placed at the start of the week and are delivered in time for the start of the following week. Now, any inventory policy whatsoever can be presented in the form of a set of rules telling what must be done if the stock is at a given level. Suppose we have the policy of ordering if and only if the stock level is one or less and that we order three units in this event. Then our policy can be presented thus:

STOCK LEVEL	ORDER
0	3
1	3
2	0
3	0
4	0

and so forth, if there are larger possible stock levels. Any policy can be presented in this way. For example, we might have such a policy as:

STOCK LEVEL	ORDER
0	3
1	2
2	1
3	0
4	0

It can be seen that this policy is of the general type of the *P*-system. The first policy example was of the general type of the *Q*-system. Note that these are not precise identifications. They are meant only as helpful approximations. We could, however, equally well have such a policy as:

STOCK LEVEL	ORDER
0	4
1	4
2	3
3	0
4	0

Why we might have such a policy is irrelevant. It is at least a possibility and such a policy fits neither of our simple types. If this policy were optimal we would have a case such as we have mentioned where a policy totally outside our simple framework is optimal. But this is a matter for subsequent discussion. Here we need only the fact that we can write any inventory policy in the form of such a table.

Granted such a formulation of the inventory policy, we can write out a matrix of the transition probabilities governing the movement of our inventory level from one time period to the next. Our time period is one week and we will define the states of the system as the inventory level at the start of the week. Let us use the first of the three policies given above as an example. Suppose our inventory level at the start of a week is zero. Then according to our policy we will order three units and they will be delivered in time for the start of the subsequent week. Since we have no stock during the current week, demand has no effect on our stock level. Therefore, if we have zero stock at the start of the week we are certain (probability is one) to have a stock of three units at the start of the following week. Suppose our inventory level at the start of a week is one. Then we will order three units and we will have either three units or four units on hand at the start of the subsequent week. We will have three units if there was a demand for at least one unit during the week and our demand distribution shows that the probability of this is .9. Otherwise (with probability of .1) we will have four units in stock at the start of the subsequent week. We can summarize this by saying that if the system is in state one there is a probability of .9 that it will be in state three in the following time period and a probability of .1 that it will be in state four.

- These are the transition probabilities leading from state one. We can proceed similarly and calculate the transition probabilities for each possible state. These can be conveniently presented in the form of a matrix:

CURRENT STATE	SUBSEQUENT STATE				
	0	1	2	3	4
0	0.0	0.0	0.0	1.0	0.0
1	0.0	0.0	0.0	0.9	0.1
2	0.5	0.4	0.1	0.0	0.0
3	0.2	0.3	0.4	0.1	0.0
4	0.0	0.2	0.3	0.4	0.1

This matrix provides a complete description of the probabilities of movement of the system from one state to another, based on the given probability distribution of demand and the specified inventory policy. It will be noted that the sum of the probabilities in any row must add up to one since, obviously, it is certain that from any state the system must proceed to some state or other.

The key question which we want to ask is: What will be the long-run behavior of the inventory level, granted the above transition or switching matrix? More specifically, what are the probabilities that the system will be in each one of its possible states? It can be proved that there are limiting probabilities which will obtain when the system is in equilibrium and which will be the probabilities that the system will be in each one of its states. To illustrate the idea involved, suppose that at some time in the future the probabilities of the system being in each one of its states are equal and, hence, equal to .2, since there are five states. Granted this, what would be the probabilities of the system being in each one of its possible states at the succeeding time interval? We can answer this question directly from the switching matrix. The probability, in the following time period, that the system would be in state zero would be simply $.5(.2) + .2(.2) = .14$. The probability that it would be in state one in the following time period would be $.4(.2) + .3(.2) + .2(.2) = .18$. The probability that it would be in state two is $.1(.2) + .4(.2) + .3(.2) = .16$. The probability that it would be in state three is $.1(.2) + .9(.2) + .1(.2) + .4(.2) = .48$. Finally, the probability that it would be in state four is $.1(.2) + .1(.2) = .04$. The sum of these probabilities is, of course, one. Given the probabilities of the system being in each one of its states for any given period we can directly, by this means, calculate the probabilities that it will be in each one of its states for any subsequent period.

If the system is going to reach equilibrium—and we can prove that it will—then the probabilities of the system being in each one of its possible states must remain constant from one period to another. In our above example the probabilities changed from one time period to the next. Thus, the probability of being in state zero in the first period was assumed to be .2 but in the subsequent period the probability of the system being in state zero was .14. This cannot happen once the system has reached equilibrium. This fact

provides straightforward means to determine the equilibrium probabilities. Let p_i represent the probability that the system is in state i . Then, at equilibrium the following set of equations would hold:

$$p_0 = .5p_2 + .2p_3$$

$$p_1 = .4p_2 + .3p_3 + .2p_4$$

$$p_2 = .1p_2 + .4p_3 + .3p_4$$

$$p_3 = p_0 + .9p_1 + .1p_3 + .4p_4$$

$$p_4 = .1p_1 + .1p_4$$

Here we have five equations in five unknowns. However, these five equations are not independent—so, as they stand, they cannot be solved. Fortunately, this causes no problem since we have another equation which we know must hold:

$$p_0 + p_1 + p_2 + p_3 + p_4 = 1$$

We can substitute this equation for any one of the set of five and solve the resulting set of five equations in five unknowns. So doing, we find the equilibrium probabilities:

$$p_0 = .17641$$

$$p_1 = .20289$$

$$p_2 = .18925$$

$$p_3 = .40891$$

$$p_4 = .02254$$

These equilibrium probabilities have the following interpretation. With the switching probabilities as given, if an observation is made at random on the system at the start of a week then the probability that it will be in the i th state is given by p_i . In other words, we can interpret these probabilities as measures of the percentage of weeks that the system will start from the given state. For example, 17.64% of the weeks will start with zero stock.

These equilibrium probabilities will establish themselves in the long run but “long run” here is often not so long. Generally, such a process will reach its equilibrium probabilities quite quickly. In any event, these probabilities describe the long-run behavior of the system and we propose to use them in order to determine the cost of the given inventory policy. The cost so determined will represent the average long-run cost of the inventory policy. For illustrative purposes, let us suppose that there is an order cost of \$10, a carrying cost of \$5 per item charged against each item in stock at the end of the week, and that each item sold represents a profit of \$20. According to our first policy we will order whenever the stock is zero or one. From our equilibrium probabilities we see that the percentage of weeks we will order is 37.93%. Therefore, our expected ordering cost per week is \$3.79. If we are in state zero we can neither sell an item nor be charged with a carrying cost for the following week. Therefore, these costs and/or profits are zero

for this state. If we are in state one we will make a \$20 profit 90% of the time and suffer a \$5 cost the remaining 10% of the time. Thus, the expected profit for this state is $.9(20) - .1(5) = \$17.50$. Proceeding similarly we can calculate the expected profit for each state, exclusive of the ordering cost:

STATE	EXPECTED PROFIT
0	\$0
1	17.50
2	25.00
3	25.00
4	20.00

The equilibrium probabilities are the probabilities that inventory will be in each one of these states at the start of the week so we can now calculate the expected profit per week of the given inventory system, remembering to subtract the expected ordering cost:

$$.1764(0) + .2029(17.50) + .1892(25) + .4089(25) + .0225(20) - 3.79 = \$15.16$$

This, then, is the expected profit per week, as a long-run average, of the stated inventory policy.

Obviously, we could similarly calculate the expected profit which would result from any specified inventory policy. As another example we can take the second policy we suggested:

STOCK LEVEL	ORDER
0	3
1	2
2	1
3	0

The reader can verify that this policy gives the equilibrium probabilities:

$$p_0 = .02439$$

$$p_1 = .30894$$

$$p_2 = .54472$$

$$p_3 = .12195$$

With these probabilities the stated policy gives an expected profit per week of \$13.30.

It is evident that we could not hope to enumerate all possible policies and evaluate each one of them in order to select that one which has the largest expected profit. Nonetheless we would like to find the optimal policy—the one that does have the largest expected profit. The next section is devoted to a method for accomplishing this on the basis of the present formulation of an inventory policy.

40. MARKOVIAN DECISION-MAKING

The method we are going to use to determine the optimal inventory policy, given the formulation of the preceding section, was developed by Ronald Howard.* The essence of this very neat procedure is that it is a variant of dynamic programming. The wary reader will immediately remember our dynamic programming example in the preceding chapter and he will recall that dynamic programming methods are algorithms for converging on a solution. They are never so pleasing as an explicit solution of an equation but thanks to these algorithms we can solve numerous problems for which we haven't the ghost of a chance to find an explicit solution. Further, he will recall that we have already pointed out that verbal descriptions of algorithms are usually far more difficult than the algorithm itself. Such is also the case here. Despite this fact, by its use we are able to solve problems which otherwise we would have no idea how to solve. With this much by way of inducement we will proceed to develop the method.

The simplest way to present this method is to use an example. We will use the example of the last section but we will make one change to start in order to simplify the discussion. The demand distribution will be as before and the profit per unit sold will be \$20 and the carrying cost per unsold unit per week will be \$5. The change is that we will ignore the ordering cost initially. With this change the first policy we considered in the preceding section has an expected profit per week of \$18.95 and the second policy has an expected profit per week of \$22.08. We want to discover the policy which has the maximum such long-run expected profit per week.

In this kind of problem we will construct our programming technique in terms of planning horizons. The planning horizon is the number of weeks we will calculate in advance in determining an optimal policy. In other words, we will start by finding the optimal policy under the assumption that we need only plan one week in advance. Then we will use the results of this calculation to find the optimal policy if we plan two weeks in advance. Then we proceed to three weeks, and so forth. Fortunately we know two major things. One is that we need not go on calculating endlessly out in the general direction of infinity. Instead, the process typically terminates quite rapidly when we use a simple criterion. Second, when we terminate we know that we have the optimal policy. We will not undertake to prove either of these statements here. The interested reader can refer to Howard's book, already cited, for the demonstrations of these two statements.

For each planning horizon we are going to be interested in determining the optimal amount of stock to order for each possible amount of stock on hand. We will need some notation to keep track of our procedure. We will let N be the number of units in stock, x the number of units we order, and we will designate by $f_i(N, x)$ the expected profit for i periods (the planning horizon) if we start with N units in stock and order x . Let us begin by calculating the

* See Ronald Howard, *Dynamic Programming and Markov Processes* (New York: John Wiley & Sons, Inc., 1960), and see the review article, "Operations Research," by Robert Dorfman, *American Economic Review*, 50 (1960), pp. 575-623.

optimal policy for $i = 1$, a planning horizon of one week. This is very easy to calculate because there is a one-week lag in delivery after order. This means that the amount we order, x , is irrelevant insofar as the expected profit for one period is concerned. Thus, we can just as well assume that $x = 0$. We could equally well assume it equalled anything else since it will have no effect on our calculations. If we have $N = 0$ we will make no profit because we have no stock to sell. Therefore, $f_1(0, x) = 0$ for all x . If we have one unit in stock we can sell it, with probability of .9, and make a profit of \$20 or we may not sell it, with probability of .1, and lose \$5 in carrying costs. The expected profit, then, is $f_1(1, x) = .9(20) - .1(5) = \$17.50$. Proceeding similarly, we calculate the expected profits resulting from each possible level of stock:

ONE PERIOD HORIZON

$$f_1(0, x) = \$ 0$$

$$f_1(1, x) = \$17.50$$

$$f_1(2, x) = \$25$$

$$f_1(3, x) = \$25$$

$$f_1(4, x) = \$20$$

$$f_1(5, x) = \$15$$

$$f_1(6, x) = \$10$$

The only reason that each one of these is constant for all x is that the lag between order and delivery happens to coincide here with the planning horizon. In all of our future iterations this will not be the case. There will be different values for each different value of x . We are interested in knowing the maximum possible return for the given planning horizon with a given amount of stock on hand, N . In this first iteration the values given in the table are constant and, hence, the maximum values for each N . To be consistent with our subsequent notation we will designate the maximum value of $f_i(N, x)$, where the function is maximized over x , as $f_i^*(N)$. Thus for the planning horizon of one week we have that $f_1^*(N) = f_1(N, x)$, as given in the preceding table.

Now we will extend our thinking to two periods, $i = 2$. Taking into account the expected profits for two periods we want to know the optimal policy to follow for each possible amount of initial inventory. This is easy—even if somewhat boring—to calculate since we can recognize that the total expected profits for two periods is the sum of two components. First is the expected profit we will make in the first of the two periods. Second is the amount of expected profit we will make if we follow the optimal policy in the second period. Consider $f_2(0, x)$. Here we have no stock to sell, and hence no profit, in the first period. Naturally, we will pick x to give us the maximum possible amount of profit in the second period. Reference to our preceding table shows that this is achieved with either x equal to two or to three. In either case we have zero expected profit in the first period and \$25 expected profit in the second period or a total of \$25 expected profit for two

periods if we start with $N = 0$. Thus, $f_2^*(0) = f_2(0, 2) = f_2(0, 3) = \$25$. Now consider $f_2(1, x)$. This is the first "typical" calculation in the sense that all of the remaining calculations will be similar to this one rather than like the special cases we have had so far. If we have one unit in stock the expected profit for the first period is \$17.50. If the one unit in stock is sold we will start the second period with x units. In either case, we will use the optimal policy for the second period. The probability of having x units at the start of the second period is directly given from the demand distribution as .9. We can therefore write the equation:

$$f_2(1, x) = \$17.50 + .9f_1^*(x) + .1f_1^*(x + 1)$$

We can now use this equation to calculate

$$f_2(1, 0) = 17.50 + .9(0) + .1(17.50) = \$19.25$$

$$f_2(1, 1) = 17.50 + .9(17.50) + .1(25) = \$35.75$$

$$f_2(1, 2) = 17.50 + .9(25) + .1(25) = \$42.50$$

$$f_2(1, 3) = 17.50 + .9(25) + .1(20) = \$42.00$$

When the successively calculated values begin to get smaller we can stop. All we are interested in is the maximum over x for the given N . In this case it is clearly $f_2^*(1) = f_2(1, 2) = \$42.50$. Interpreted in words, if we calculate only over two periods and if we start the first period with one unit in stock, then the optimal policy is to order two units. The expected profit for this policy is \$42.50.

We must continue similarly. We need $f_2(2, x)$. Starting with two units in stock in the first of the two periods we will make an expected profit of \$25. If we order x units then we will have either x , $(x + 1)$, or $(x + 2)$ units in stock at the start of the second period. We can determine the probabilities of these three possibilities directly from the demand distribution and we have the equation:

$$f_2(2, x) = 25 + .5f_1^*(x) + .4f_1^*(x + 1) + .1f_1^*(x + 2)$$

Using this equation we can calculate:

$$f_2(2, 0) = 25 + .5(0) + .4(17.50) + .1(25) = \$34.50$$

$$f_2(2, 1) = \text{similarly} = \$46.25$$

$$f_2(2, 2) = = \$49.50$$

$$f_2(2, 3) = = \$47.00$$

Obviously, we have here $f_2^*(2) = f_2(2, 2) = \$49.50$. In words, if we calculate only over two periods and if we start the first period with two units in stock then the best policy is to order two units.

We proceed to $f_2(3, x)$ for which a straightforward repetition of our reasoning leads to the equation:

$$f_2(3, x) = 25 + .2f_1^*(x) + .3f_1^*(x + 1) + .4f_1^*(x + 2) + .1f_1^*(x + 3)$$

From this equation we calculate:

$$f_2(3, 0) = \$42.75$$

$$f_2(3, 1) = \$48.00$$

$$f_2(3, 2) = \$47.00$$

Evidently, $f_2^*(3) = f_2(3, 1) = \$48.00$ with the usual verbal translation.

For $f_2(4, x)$ we have the equation:

$$f_2(4, x) = 20 + .2f_1^*(x + 1) + .3f_1^*(x + 2) + .4f_1^*(x + 3) + .1f_1^*(x + 4)$$

Using this we calculate:

$$f_2(4, 0) = \$43$$

$$f_2(4, 1) = \$42$$

Thus, $f_2^*(4) = f_2(4, 0) = \$43$. And, finally, we calculate by similar procedures that $f_2^*(5) = f_2(5, 0) = \$37$. We have no need to calculate further for f_2 since it is clear that the optimal policy for an N greater than five will also be to order zero units.

For the purposes of subsequent iterations all of these calculations are summarized by giving the f_2^* values:

N	$f_2^*(N)$	$x_{\max}$
0	\$25	2 or 3
1	42.50	2
2	49.50	2
3	48	1
4	43	0
5	37	0

These are the optimal policies for a horizon of two periods and they are all we need as we proceed to a consideration of a horizon of three periods. We give the values of x which produce the f_2^* values for future ease of reference.

When we want to calculate an optimal policy on the basis of a three-period horizon we can simply calculate the amount we will make in the first period plus the expected value of the use of an optimal policy in the two subsequent periods. Thus, $f_3(0, x)$ provides no stock for sale in the first period and an optimal policy for the two succeeding periods. Inspection of our table of $f_2^*(N)$ values discloses that the optimal profit will result if we order two units and follow an optimal policy thereafter. Thus, $f_3^*(0) = \$49.50$. Suppose we start the first period with one unit in stock. Then our expected return in the first period is \$17.50. If we order x units, then with probability .9 we will start the second period with x units, and with probability .1 we will start the second period with $(x + 1)$ units. This gives us the equation:

$$f_3(1, x) = 17.50 + .9f_2^*(x) + .1f_2^*(x + 1)$$

Thus, we calculate:

$$f_3(1, 0) = 17.50 + .9(25) + .1(42.50) = \$44.25$$

$$f_3(1, 1) = 17.50 + .9(42.50) + .1(49.50) = \$60.70$$

$$f_3(1, 2) = \text{similarly} = \$66.85$$

$$f_3(1, 3) = = \$65.00$$

Clearly, $f_3^*(1) = f_3(1, 2) = \$66.85$.

The next equation is

$$f_3(2, x) = 25 + .5f_2^*(x) + .4f_2^*(x + 1) + .1f_2^*(x + 2)$$

and we calculate:

$$f_3(2, 0) = \$59.45$$

$$f_3(2, 1) = \$70.85$$

$$f_3(2, 2) = \$73.25$$

$$f_3(2, 3) = \$69.90$$

Thus, $f_3^*(2) = f_3(2, 2) = \$73.25$.

Proceeding in this way we can complete the table representing this iteration:

N	$f_3^*(N)$	$x_{\max}$
0	\$49.50	2
1	66.85	2
2	73.25	2
3	71.85	1
4	66.85	0
5	60.20	0

The calculations for f_4^* proceed in the same way. As an example, the equation for $f_4(2, x)$ is determined exactly as previously and is

$$f_4(2, x) = 25 + .5f_3^*(x) + .4f_3^*(x + 1) + .1f_3^*(x + 2)$$

Calculating all of the f_4^* values gives

N	$f_4^*(N)$	$x_{\max}$
0	\$73.25	2
1	90.60	2
2	97.04	2
3	95.77	1
4	90.77	0

We could continue this process indefinitely but, fortunately, we don't have to do so. Remembering that our primary purpose has been to determine the optimal policies for various possible levels of initial stock and for steadily increasing planning horizons, let us sort out our conclusions so far in a different way. What have we discovered about the policies that are

optimal as we have extended the planning horizon? The following table provides the answer to this question:

INITIAL STOCK	PLANNING HORIZON			
	1	2	3	4
0	0	2 or 3	2	2
1	0	2	2	2
2	0	2	2	2
3	0	1	1	1
4	0	0	0	0

In the body of the table are given the values of $x_{\max}$ and these represent the optimal policies for the given combination of initial stock and planning horizon. Now, we are able to stop the calculations as soon as two successive iterations have produced the same optimal policy. This has happened in our case for planning horizons of three and four periods. Therefore we have completed our analysis. The optimal policy is the one given under the planning horizons on three and four periods. This policy is to order two units if the stock on hand is 0, 1, or 2 units, to order one unit if the stock on hand is three units and to order nothing otherwise. To find the long-run expected profit from this policy we must determine the switching matrix and proceed as in the last section. It will be discovered that the probability of zero stock, at equilibrium, is zero. The equilibrium probabilities are

$$p_1 = .09259$$

$$p_2 = .44444$$

$$p_3 = .37223$$

$$p_4 = .09074$$

Using these probabilities we can calculate the expected long-run profit per week as \$23.85. This is the largest such expected profit which can be achieved with any inventory policy. Thus, the problem we started with has been solved. It will be seen that the optimal policy bears no resemblance to either of the kinds of policies we have previously considered. For the simple problem we have considered, this is the solution to the general question—the policy we have discovered is the optimum over all possible inventory policies.

We did not include the ordering cost in the above example in order to make the reasoning as easy as possible. It is, of course, a straightforward matter to include the ordering cost. Since the understanding of such an algorithm as this one is not complete until the method is tried out the reader may want to try including the ordering cost himself. To do so it is only necessary to repeat the form of the above calculations but subtracting \$10 from the expected profit each time an order is included. One precaution may be noted. In our calculations above we only needed to calculate each $f_i(N, x)$ for $x = 0, 1, 2, \dots$, until $f_i(N, x + 1)$ was less than $f_i(N, x)$, at which point we could stop. We can use the same rule here with one proviso. This is that we must treat $f_i(N, 0)$ as a separate case. In other words, the fact that

$f_i(N, 1)$ may be less than $f_i(N, 0)$ is not sufficient reason to stop the calculations. Otherwise, the procedure is exactly the same as before. It will be found that including the order cost of \$10 results in the following optimal policy:

STOCK	ORDER
0	4
1	4
2	3
3	0
4	0

It will be seen that this policy is not one of the simple kinds that we considered earlier in this chapter. Constructing the switching matrix for this policy and calculating the equilibrium probabilities gives:

$$p_0 = .05229$$

$$p_1 = .13943$$

$$p_2 = .20370$$

$$p_3 = .26144$$

$$p_4 = .30501$$

$$p_5 = .03813$$

In accordance with our inventory policy we will make an order in 39.54% of the weeks. The expected profit per week from this policy is \$16.79. No other policy has a larger expected profit than this one. This value can be compared to the long-run expected profits of the policies we calculated in the last section.

The methods of this section, then, provide means for determining the true optimal policy over all possible policies. By utilizing computers we could apply this method for discrete probability distributions of demand where a large number of levels of demand are possible. Furthermore, we can mathematically investigate the behavior of continuous probability distributions of demand under analogous procedures. However, unless we consider only rather sharply circumscribed demand distributions we cannot achieve any significant general results in this latter case. Nonetheless, this procedure is a major step in the direction of having methods for determining the true optimal inventory policy.

41. OPTIMAL SELLING POLICY WITH FLUCTUATING PRICES

Since Section 40 was devoted to a discussion of one of the uses of dynamic programming it will be convenient to devote this section to another use of these valuable methods. This example is considerably simpler than the preceding one. The problem we are going to consider is an interesting one and variants of it frequently arise in practice. It is closely related to inventory problems and this will be our only occasion to discuss it. The problem has to do with the discovery of optimal policies when faced with

fluctuating prices.* It can equally well be stated in terms of the problem of buying some item or commodity when faced with fluctuating market prices or in terms of selling an existing inventory under similar circumstances. The same procedures suffice to solve either variant of the problem. We will state, and solve, the problem in terms of selling. Suppose we already have an inventory which we intend to sell at some time during a demand season. The price of the item in inventory fluctuates during the season and we would like to sell, naturally, for the largest price we can get. Such problems arise frequently, for example, in the case of commodities such as soybeans, cocoa, or eggwhites. What is the optimal policy we should follow in order to ensure that we get the largest possible price for our inventory of the item in question? The other variant of the problem, buying the item would require us to attempt to achieve the minimum possible price. The procedure will be exactly the same as what follows, with some obvious modifications.

In order to determine an optimal selling policy we need to have some information about the price fluctuations over the demand season. In accordance with the underlying theme of this chapter we will assume that we have knowledge of the probability distribution of prices. Such information would ordinarily come from analysis of historical data concerning the price fluctuations of previous seasons. However, there is a question of the time basis on which these probability distributions of price should be based. At one extreme we might consider simply determining the probability distribution of price over the entire demand season. Apart from questions of meaningfulness, this kind of presentation would be of no help to us whatsoever. The whole point of the problem is that there is a time period during which we can sell our inventory. We must have information about the price probability distribution at different points in time in order to have any hope of finding an optimal selling policy in terms of the entire time period. At the other extreme we might consider determining the price probability distribution for each day, say, during the entire time period. There would be nothing wrong with this procedure insofar as usefulness is concerned but practical exigencies would generally prevent its accomplishment. The major such practical difficulty would be the lack of a sufficient amount of historical data to permit the breakdown into a daily probability distribution of price. But even if there were sufficient data we would not ordinarily be satisfied with such a breakdown. We intend to use probability distributions which obtained in the recent historical past as predictions of what is going to happen in the immediate future. Our confidence in such a procedure depends on the existence of some kind of stability in the process so described. In a case such as the one we are considering, any daily breakdown would be quite erratic and would not inspire confidence in the continuation of the pattern into the future. Therefore, what we will generally want will be some intermediate breakdown—long enough to show some stability and short enough so that several of the periods over which the probability distributions are given will be

* See William T. Morris, "Some Analyses of Purchasing Policy", in *Management Science*, Vol. V (1959), pp. 443-452.

included in the demand season. No set rules can be given to determine the best way in which to accomplish this but familiarity with the price fluctuations of the item in question will usually help to indicate the time period which should be used in determining the probability distributions of price.

Let us then suppose that for the item we are considering there is a five-month selling season and that we have probability distributions of price for each month of the season. We can conveniently present this information in the form of a table:

PRICE	MONTH				
	1	2	3	4	5
\$4	.10	.05	.05	.05	.05
5	.10	.10	.15	.05	.25
6	.20	.15	.30	.35	.30
7	.30	.25	.15	.25	.20
8	.15	.20	.15	.15	.15
9	.10	.15	.10	.10	.05
10	.05	.10	.10	.05	.00

From these probability distributions we can calculate the expected prices for each of the five months:

MONTH	EXPECTED PRICE
1	\$6.80
2	7.30
3	6.90
4	6.90
5	6.30

The question is: Faced with these probability distributions of price over the demand season, what would be our selling policy? We would expect that such a policy would consist of rules telling us that if in a given month the actual price exceeded a certain amount, we should sell. How can we determine such rules?

The method we will use is another variant of dynamic programming. We will reason from the last month of the season backwards in an iterative procedure. Suppose that we come to the last month of the season, *M5*, and have not yet sold our inventory. Then we will have no choice but to sell it and the expected price we will get is the expected price for that month, \$6.30. Since we can do this well in the last month it follows that we will not sell in the penultimate month, *M4*, unless we can do better than this. In other words, we will only sell in *M4* if the price is greater than \$6.30. Using the rather gross price intervals we have chosen for illustrative purposes—and we could equally well have used much finer intervals if we could have determined the probability distributions for them—this means that we should sell in *M4* only if the price is \$7 or more. Looking at our probability distribution for *M4* we see that the probability that the price will be \$7 or more is .55. This means that the probability, in terms of the last two months only, that we will sell in *M4* is .55 and the probability that we will sell in

$M5$ is therefore .45. Taking into account the individual prices at which we might sell in $M4$ we can calculate the expected price of the policy of selling in $M4$ only if the price is greater than or equal to \$7 and selling in $M5$ otherwise:

$$\begin{aligned}\text{Expected price} &= .25(7) + .15(8) + .10(9) + .05(10) + .45(6.30) \\ &= \$7.185\end{aligned}$$

This is the expected price of our policy for the last two months.

Clearly, since we can do this well in the last two months we will not sell in the antepenultimate month, $M3$ unless we can do better than this. Therefore, we will only sell in $M3$ if the price is more than \$7.185, or greater than or equal to \$8. From our table we see that the probability of this is .35 and, therefore, the probability that we will sell in one of the last two months, at an expected price of \$7.185, is, .65. Again we can calculate the expected price for the policy, including $M3$:

$$\text{Expected price} = .15(8) + .10(9) + .10(10) + .65(7.185) = \$7.770$$

We will sell in $M2$ only if the price is greater than this expected price for $M3$. The probability of this is the probability that the price in $M2$ will be \$8 or more. This probability is .45 so the probability that we will sell in one of the last three months is .55. We can calculate the expected price resulting from this policy for $M2$:

$$\text{Expected price} = .20(8) + .15(9) + .10(10) + .55(7.77) = \$8.224.$$

Finally, we will only sell in $M1$ if we can do better than this, meaning that the price in $M1$ is greater than or equal to \$9. The expected price for this policy will be

$$\text{Expected price} = .10(9) + .05(10) + .85(8.224) = \$8.390$$

This is the expected price which we can achieve by following the optimal policy we have deduced. The optimal policy can be summarized:

$M1$: Sell only if the price is greater than or equal to \$9.

$M2$: Sell only if the price is greater than or equal to \$8.

$M3$: Sell only if the price is greater than or equal to \$8.

$M4$: Sell only if the price is greater than or equal to \$7.

$M5$: Sell.

The expected price we can achieve by this policy is \$8.39.

One immediate question concerning this policy arises. We have not considered the possibility of partial sales in different months. We have tacitly assumed that the whole inventory will either be sold at once or not at all. This seems contrary to the whole philosophy underlying hedging and, therefore, the realism of our analysis can be disputed. We will answer this criticism in two steps. First, we will prove that the policy of selling all or none is optimal. Let $E(n)$ stand for the expected selling price per unit over n periods. Suppose that we have a total quantity, Q , to sell and that we

consider selling some partial amount, q , when there are n periods remaining. Let the price at which we are going to sell these q units be P per unit. Then the expected selling price under this policy, $E(n)$, will be

$$E(n) = \frac{qP + (Q - q)E(n - 1)}{Q}$$

where the right-hand side simply gives the total income which will result from the policy, q units sold at P and $(Q - q)$ units sold at the expected selling price for the remaining $(n - 1)$ periods, divided by the total quantity sold. This can be rewritten as

$$E(n) = \frac{q}{Q} [P - E(n - 1)] + E(n - 1)$$

Now, if P is less than $E(n - 1)$ it follows that $E(n)$ will be less than $E(n - 1)$ unless we put $q = 0$, which means sell none. If P is greater than $E(n - 1)$ we will maximize the advantage of $E(n)$ over $E(n - 1)$ by increasing q to the maximum possible amount, which means to put $q = Q$ and to sell all of the stock. Therefore, in order to achieve the largest possible expected price per unit, and the largest total income, we need not consider partial sales.

For the second step of our answer we will note that the proof we have just given of the optimality of the "all or none" selling policy seems to be a bit of mathematical legerdemain because the fact still remains that in practice businessmen frequently hedge. Our proof would indicate that they shouldn't hedge. Are we to believe that the businessmen who do hedge are wrong? The answer is no. We have proved only that hedging is not an optimal policy if the decision-maker is governed by expected values. If some other criterion is involved our argument may not hold. Specifically, if the businessman is governed by the Savage criterion we discussed in Chapter 3 then it can be shown that hedging is sometimes the best policy. We will not tarry to demonstrate this here but merely note that our analysis is correct unless the Savage criterion is being used by the decision-maker.

42. VARIABLE LAGS BETWEEN ORDER AND DELIVERY: THE QUEUING MODEL

As a final example for this chapter we will consider a problem which frequently arises and which often causes serious analytical difficulties. This is the problem of inventory policies where there is random variation in the length of time it takes to get an order filled. This problem arises often and it can equally well occur in connection with any kind of item. However, for many items the variation in delivery time is so small in relation to the average delivery time that it can be ignored. One simply takes the maximum delivery time instead of the average. The kinds of items for which this simplification cannot reasonably be used are often characterized by another set of common features as well. Namely, that these kinds of items are frequently of the sort that are made to order and, hence, often on order. Typically such items are very expensive and, furthermore, the out-of-stock cost on such an item

is likely to be very high. Examples of such items might be large compressors, or large motors, specialized drilling tools, and so forth. It is the inventory problem for items such as these that we want to consider in this section.

The major features of such inventory problems from an analytical point of view are that carrying costs are very high, out-of-stock costs are very high, and demand is likely to be very low. As a result of the first two points the optimal inventory policy will usually be obtained by basing the total cost equation only on the carrying and out-of-stock costs. The order cost, in particular, is so small relatively that it can be ignored. The fact that demand is low will typically result in the fact that the probability distribution of demand will be a Poisson distribution. This fact has no particular significance if the lead time is constant but it does when there is a probability distribution for the lead time.

Let us first consider the inventory problem under the assumption that there is a constant lag between order and delivery. As an example we will assume an item which costs \$4000 and for which the time between order and delivery is four months. We will assume that the carrying cost is 18% per year and that the out-of-stock cost for this item is a fixed \$50,000. Finally, we will assume that the average demand for the item is 1.5 units per month. With regard to the distinguishing features we mentioned above, this item is more or less typical. Since the order cost is inconsequential relative to the carrying cost and the out-of-stock cost we would usually ignore it entirely. The optimal inventory policy for this item would be to determine the necessary reserve stock and then to order one item each time there is a demand for one item. Thus, the analytical problem is simply to determine the optimal reserve stock. This can be done immediately by the methods we developed in analyzing static inventory problems under risk. When the average demand is so low it will often be the case that the demand distribution is a Poisson distribution. This is the assumption we will make here. The Poisson distribution is completely determined when the mean of the distribution is known. Further, when a random variable has a Poisson distribution over one time period then it will have a determined Poisson distribution over any greater time period. Since we know the average monthly demand we can determine the demand distribution for the four month period between order and delivery. This distribution is given by the equation:

$$P(z = i) = \frac{e^{-6}(6)^i}{i!}$$

The left-hand side of this equation reads: the probability that demand will equal i units. The equation is given for a four-month period so it gives us immediately the probability of any specific demand during the lead time for the item. Expressing our total cost equation on the basis of the lag period between order and delivery we have:

$$T.C_r = 240x + 50,000 \sum_{i=x+1}^{\infty} \frac{e^{-6}(6)^i}{i!}$$

The 240 in the equation comes from $(\$4000)(.18)(1/3) = \$240$, the carrying

cost per unit for four months. This equation can be directly solved by reference to a table of the Poisson distribution. We want to find that value of x for which $T.C_{x+1}$ is greater than $T.C_x$. This is clearly the minimizing value of x . Now, $T.C_{x+1}$ will be greater than $T.C_x$ when

$$\frac{240}{50,000} > (e^{-6}) \frac{6^{x+1}}{(x+1)!}$$

Reference to tables shows that this is first true for our distribution when $x = 13$. This, then, is the optimal amount of stock to be carried. The procedure is, as already stated, that an order is placed for one unit each time there is demand for one unit.

Thus, the inventory problem is easy to solve when the delivery time is a fixed interval. . . What happens when there is variability in the delivery time? Obviously, it then becomes necessary to maintain some additional number of units in stock in order to provide protection against the variability in delivery time. The question, of course, is how many. The answer to this question is most easily obtained by having recourse to a different kind of operations research model, the queuing model. Queuing theory is a large and intricate subject in itself. In Part 2 we will again have occasion to utilize queuing theory as an important contributor to our systems analysis. However, for our present purposes, we can quickly summarize the information which we need in order to utilize a queuing model to solve this inventory problem. Queuing theory deals with the problem of providing adequate service facilities to handle an arriving stream of things, or people, requiring service of some kind. The theory is very general and is equally useful if telephone calls requiring telephone wires, customers in a restaurant requiring tables, airplanes requiring landing strips, broken machines requiring repair, or any of an incredible number of other kinds of servicing needs are under consideration. Generally, queuing models are classified in accordance with the probability distribution of arrivals and the probability distribution of servicing times. In addition, the number of servicing units—called channels—is of key importance. The analogy of such models to our inventory problem may not seem immediately obvious. However, we can easily point out the similarities. First, many queuing models assume a Poisson distribution of arrivals over time. We have already indicated that for items of low demand the Poisson distribution is a very likely candidate for being the demand distribution. Second, when demand occurs for a unit it is clear that we can consider this to be demand for a service—namely, the service of having a unit in stock to fill the demand. Further, the number of units in stock is exactly analogous to the number of service channels. Queuing models are concerned with such things as the probability of something arriving for service and having to wait because all of the channels are occupied. In our inventory problem this would occur if we were out of stock. Finally, the queuing models are constructed to take explicit account of probability distributions of service times. Our service times are the times it takes to replace a unit in stock and precisely what we want to do is to take account of variability in these times. Therefore, it appears that queuing models may be just what we need.

The particular queuing model which we propose to use assumes an exponential distribution of service times. In order for this model to be completely applicable to our problem we would have to have an exponential distribution of times between order and delivery. As a matter of fact, this is not a bad approximation in many cases and we will assume that it holds true of the item we have under consideration. It should be noted, however, that it is possible to extend the queuing models to take into account other possible distributions of service times if this is necessary.* Adhering to our intention to avoid further development of queuing theory in this context we intend to simply cite the relevant equation and we refer the interested reader to the literature for its derivation.† The specific queuing model we will use is called—in the queuing literature—an infinite source, multiple channel model with Poisson arrivals and exponential holding times. For this model the probability of an out-of-stock can be written thus:

$$P_o = \frac{1}{1 + \frac{(x-1)!(x-6)}{6^x} \left[1 + 6 + \frac{6^2}{2!} + \dots + \frac{6^{x-1}}{(x-1)!} \right]}$$

The relevant total cost equation is, then:

$$T.C_x = 240x + 50,000P_o$$

This is not a particularly pleasant kind of equation to minimize unless special tables of P_o are available. However, it can be done without them. The expression in brackets in the denomination of P_o can be quickly determined from tables of the cumulative Poisson distribution. This means that it is not too difficult to calculate P_o explicitly for various values of x . We want that x for which $T.C_{x+1} > T.C_x$ or

$$.0048 > P_o(x) - P_o(x+1)$$

Trying some values we discover that

$$P_o(15) = .00148$$

$$P_o(14) = .00389$$

$$P_o(13) = .00964$$

Therefore, our inequality first holds for $x = 14$ and this is the amount of stock to be carried.

Comparison with our previous result shows that we need only carry one additional unit in stock in order to achieve satisfactory protection against the variability in delivery time. Naturally, the additional amount required depends on the costs and the specific data of the given problem. Sometimes a very much greater amount of stock is required because of the variability in delivery times. Other times, as here, the effect of this variability is slight.

* See Fig. 9-3.

† See the Bibliography at the end of this book for literature dealing with queuing theory.

In any event, it is possible to handle this kind of problem by using the methods and models of queuing theory as illustrated in our example.

PROBLEMS

1. You are given the demand distribution for one week:

DEMAND	PROBABILITY
0	.2
1	.5
2	.2
3	.1

What is the demand distribution for four weeks?

2. (a) Weekly demand for an item is normally distributed with mean equal to 100 and standard deviation equal to 8. Ordering cost is \$10 and carrying cost is 12% per year. The price per item is \$12. There is a two-week lag between order and delivery. Assume there is a fixed understock cost of \$200. Using the approximate method of analysis, find the optimal Q -system for this item.

(b) What is the total cost for this system?

(c) Use the exact method of analysis to find the optimal Q -system.

3. (a) For the item of Problem 2, use the approximate method of analysis to find the optimal P -system for the item. Make the assumption that the reserve stock is carried continuously in calculating carrying costs.

(b) What is the total cost for this system?

(c) Use the exact method of analysis to find the optimal P -system for this item.

4. (a) An item has the demand distribution of Problem 1. Assume that an order placed on the first day of one week is delivered in time for sale on the first day of the following week. The item is reviewed on the first day of each week and an order is placed the same day according to the following schedule:

STOCK ON HAND	ORDER
0	3
1	3
2	1
3	0
4	0

What is the switching matrix for this item?

(b) What are the equilibrium probabilities for the various states?

(c) There is an order cost of \$10 and a \$6 carrying cost is charged against each item in stock at the end of the week. Each item sold produces a profit of \$25. What is the long-run expected profit per week of the above policy?

(d) Use the method of Section 40 to determine the optimal policy for this item.

5. An item has a six-month selling period with the following probability distributions of selling prices in each month:

PRICE	MONTH					
	1	2	3	4	5	6
\$12	.20	.15	.10	.05	.05	.05
13	.15	.15	.10	.10	.25	.25
14	.10	.10	.30	.15	.35	.30
15	.05	.10	.20	.25	.20	.25
16	.10	.15	.15	.20	.10	.10
17	.15	.20	.10	.15	.05	.05
18	.25	.15	.05	.10	.00	.00

What is the optimal selling policy?

dynamic inventory problems under uncertainty

43. GENERAL CHARACTERISTICS

Dynamic inventory problems under uncertainty are characterized by the possibility of placing more than one order and by the fact that we do not know the probability distribution of demand. We will not repeat the discussion of Chapter 3 here but we will recall the fact that we are including under uncertainty not only complete ignorance about the probability distribution of demand but also partial information about it. It is this latter case which is far more practical and important. We are rarely totally ignorant and it is equally rare that we have complete information about anything so recondite as a demand probability distribution. This chapter will consider solely the intermediate cases—where we have some information but not enough to completely characterize the demand distribution. We will not consider the case of complete ignorance here at all. There is nothing more to say about this case than we said in the chapter dealing with static inventory problems under uncertainty and we will not repeat that discussion. The reader is referred to that chapter for a discussion of the different criteria which may be used in the case of total ignorance.

As we noted in the preceding chapter, dynamic inventory problems under risk or under uncertainty constitute the overwhelming majority of everyday inventory problems. The question of which of the two categories a specific inventory problem falls in is solely a question of what information we have about the demand distribution. In the authors' experience it is far more frequent that we do not have complete information than that we do. If this is true it follows that the most frequent kind of inventory problem is precisely the subject of this chapter.

In our discussions of specific kinds of this type of inventory problem we will use the same distribution-free approach we used in the corresponding chapter dealing with static problems. The same inequalities which played a major role in our analyses in that chapter will be used here. Since we do not want to duplicate previous arguments we will refer the reader to that chapter for a discussion of these inequalities and a proof of some of them. For our present purposes we will assume a knowledge of these basic inequalities.

One point is worth re-emphasis. In considering the results of methods based on partial information it should be borne in mind that it is not reasonable to compare them to the results of analyses based on complete information and then to conclude that they are not good enough because they are considerably less efficient. It is obviously the case that we can do better if we have complete information. However, if we really only have partial information then calculations based on the assumption of complete information are meaningless. If we fully utilize all of the information we have then the resulting inventory system is the best it is possible to achieve—and that should end the discussion. If the resulting policy is not satisfactory then a serious effort must be undertaken to get more information so that a more precise solution to the given inventory problem can be achieved. These remarks do not mean that one should not compare the results of analyses based on partial information with those based on total information. Such comparisons are always useful in determining the worth to the decision-maker of additional information. We are only concerned with defending the methods based on partial information against illogical criticism.

Our general approach in this chapter will be to consider the two simple kinds of inventory systems discussed in the last chapter and see what can be done with them when we have only partial information about the demand distribution. In so doing we must make the same assumptions that we made in the last chapter. First, we will assume that the demand distribution, whatever it is, is the same in each period and, in particular, that demand in one period is independent of the demand in the preceding periods. Second, we recognize that the problem of determining the true optimal inventory policy over all possible policies is a very difficult one and we will be satisfied to find the optimal policy within some subclass of all possible policies. Obviously, we cannot hope to develop any distribution-free analysis in analogy to the switching matrix formulation of the last chapter because this approach requires complete information about the demand distribution. Finally, we will try to analyze some important special problems in terms of having only partial information.

44. THE MOMENTS OF CONVOLUTIONS

As in the preceding chapter we will need to consider the convolutions of the demand distribution. The reason for this is that we will usually know the demand distribution for some given interval of time and we need to know it for various multiples of this basic time period. The procedure, which

obtains the demand distribution for the longer period, requires taking the convolutions of the original basic distribution. This problem may appear to have originated with the needs of our analysis of dynamic inventory problems. However, this is not really true. The problem is equally likely to be involved in the static analyses. We did not introduce convolutions in the static analyses but this was purely for pedagogic reasons. Even in the static analysis we might have to take the convolutions of a given demand distribution. This would occur, for example, if we knew the demand distribution for a one-month period and had to place one order to cover a one-year period. In this event we would need to convolute the one-month distribution in order to determine the one-year demand distribution.

The inequalities which we used for our distribution-free analyses of static inventory problems, and which we are going to use again for the dynamic problems, are based on various moments of the demand distribution. For the analyses of the preceding chapter it was necessary to assume complete knowledge of the demand probability distribution and it was assumed that this probability distribution was the same for every equal time period. We will assume, in our distribution-free analyses, that the probability distributions of demand for equal time periods are equal but we will not assume that we have complete information about the distribution. We cannot, therefore, convolute the basic probability distribution. However, we will need to know how to reason from the moments of the basic distribution to the moments of the convolutions of this distribution if we hope to use our inequalities on the various convolutions. Fortunately, the relationships between the moments of the convolutions and those of the basic distribution are quite simple, at least for the lower order moments which we will be using. Let us suppose that the demand distribution is known for some fixed time period. Then the i th convolution of the demand distribution, $f^i(y) dy$, will be the demand distribution for i time periods. Let the moments of the basic distribution be $\bar{z}$, s^2 , m_3 , and m_4 . Then it can be shown that the corresponding moments of the i th convolution will be

$$\bar{z}_i = i\bar{z}$$

$$s_i^2 = is^2$$

$$m_{3,i} = im_3$$

$$m_{4,i} = im_4 + 3i(i-1)s^4$$

The equations for the higher order moments can also be written out but they become considerably more complicated. The measures of skewness and kurtosis, γ_1 and γ_2 , may also be mentioned with regard to convolutions. In the Appendix we derive the fact that

$$\gamma_{1,i} = \frac{\gamma_1}{\sqrt{i}} \quad \text{and} \quad \gamma_{2,i} = \frac{\gamma_2}{i s^4}$$

Not all of the kinds of information we might reasonably have concerning the demand distribution can be expressed in terms of the moments. In our static analyses we often utilized the information that the demand distribution

was symmetrical. Extending this to the convolutions of the distribution we have the fact that if the original distribution is symmetrical then so will be all of the convolutions of the distribution. Similarly, if the original distribution was unimodal then so will be all of the distribution's convolutions. No general expression can be given which would relate the modes of the various convolutions to the characteristics of the original distribution. However, for the Pearsonian family of distributions if γ_1 and γ_2 are small we can derive the following simple approximation:

$${}_iM_o = im_i - \frac{1}{2}\gamma_1s$$

We will not present here any methods for reasoning from the moments about the mode of the original distribution to the same kind of moments for the convolutions of the distribution.

These are the relationships we will need to use in extending our distribution-free analyses to dynamic inventory problems. Exactly as in the case of the static analyses, we would like to have methods for utilizing whatever information we have concerning the demand distribution. The inequalities we are going to use seem the most expeditious means for accomplishing this purpose.

45. *Q*-SYSTEMS FOR INVENTORY CONTROL

It will be remembered that the *Q*-system of inventory control calls for a fixed order size and a variable order period. In terms of the analysis of this kind of system the basic point is that there is need only for reserve stock to protect against fluctuations in demand during the lag time between order and delivery. In our original discussion of the *Q*-system we first used some approximate methods for determining the optimal *Q*-system and then formulated an exact equation, which we found gave very nearly the same solution as the approximate solution. We will follow the same procedure here as we proceed to develop some distribution-free analyses of *Q*-systems.

The approximate method of analysis consists of determining the expected order period for the given item from the optimal lot-size equation and then determining separately the reserve stock for the lead time. We will use as an example the same basic data we used in the preceding chapter. For this item on a weekly basis we have $\bar{x} = 50$ and $s = 5$. The price per unit is \$5, the carrying cost is 12% per year, and the order cost is \$10. The lead time is three weeks. We are not assuming here, of course, that demand is normally distributed since such an assumption would preclude the need for distribution-free analysis. For this item the optimal lot-size equation gives an optimal order size of $x_o = 294.4$, as calculated in Section 37. This is equivalent to an average order period of $294.4/50 = 5.89$ weeks. Our procedure is to determine separately the optimal reserve stock as protection against demand fluctuations in the three-week lead time period. As a reasonable approximation we can assume that this reserve stock will be carried constantly in stock and, therefore, the carrying costs for it can immediately be expressed

on a yearly basis. We will be balancing these carrying costs against the cost of being out-of-stock. Alternatively, if the out-of-stock cost is difficult to measure we may be satisfied to use some performance standard established—generally intuitively—by management. There are, thus, three cases which can be considered:

- (1) Performance standard set by management.
- (2) Fixed out-of-stock cost regardless of the number of units short.
- (3) Per unit out-of-stock cost.

Let us begin with the first case and assume, as in Section 37, that management has set the standard that the level of stock should be sufficient to ensure that there is not an out-of-stock more than one time in a thousand lead time periods. Using such a performance standard does not permit the evaluation of the total cost and we can directly utilize our inequalities to determine the necessary reserve stock level. Thus, if all we know is the mean and standard deviation of the demand distribution on a weekly basis then we know that the demand distribution on a basis of a three-week period will have a mean of 150 and a standard deviation of 8.66. Using Tchebycheff's inequality we know that

$$P(y - 150 \geq 8.66k) \leq \frac{1}{k^2}$$

We want this probability to be less than or equal to .001. Therefore, $k = \sqrt{1000} = 31.62$ and we need to carry $8.66(31.62) = 273.8$ units in reserve. Now, this is undoubtedly an absurdly high level and we will return to some comments concerning it shortly. First, however, we will point out that further information about the demand distribution would enable us to find a more precise level for the reserve stock. If we know the demand distribution is symmetrical we can immediately set $1/2k^2 = .001$ and find $k = 22.36$ with a resulting reserve stock of 193.6. Suppose we know that the measure of kurtosis, γ_2 , for our demand distribution equals zero. Then we have immediately

$$P[(y - \bar{z}) \geq ks_i] \leq \frac{\gamma_{2,i} + 2}{(k^2 - 1)^2 + \gamma_{2,i} + 2} = .001$$

With $\gamma_2 = 0$ and, therefore, $\gamma_{2,i} = 0$ we calculate that $k = 6.76$ or a reserve stock of 58.5. Similarly, we can utilize various other inequalities if we have other kinds of information about the demand distribution.

Now, let us return to the remark we made in the preceding paragraph about the absurdity of the reserve stock determined from the Tchebycheff inequality. Such a remark seems in flagrant contradiction to our repeated emphasis on the fact that one can do no better than to utilize all the information one has and that it does little good to bemoan the fact that one doesn't have more information if there is no way to get it. However, this contradiction is more apparent than real because we suggest absurdity on the basis of the imputed cost of being out-of-stock, not on the mere fact that the reserve stock was high. To illustrate the argument involved let us turn our attention

to the second case listed above, where there is a fixed out-of-stock cost. To analyze this case on a distribution-free basis, and still using the approximate analysis, we would set up the total cost equation:

$$T.C_r = rcC_c + \frac{52}{t} KP(y > r)$$

where t is the weeks per order period, $52/t$ is the order periods per year, K is the fixed cost of being out-of-stock, and $P(y > r)$ is the probability of being out-of-stock in the lead time interval. Using the Tchebycheff inequality we would write:

$$T.C_r \leq ks_i c C_c + \frac{52}{t} K \frac{1}{k^2}$$

The expression on the right-hand side of this equation is the maximum possible total cost for a given k and we minimize this maximum with regard to k :

$$\frac{dT.C_r}{dk} = s_i c C_c - \frac{104K}{tk^3} = 0$$

or

$$k = \sqrt[3]{\frac{104K}{ts_i c C_c}}$$

From this minimizing expression for k it can be seen that k is a function of K and, hence, that K is a function of k in the sense that a specified value of k , the other quantities in the expression being known, imputes a value of K . For our first example above we found $k = 31.62$. From this we calculate that

$$K = \frac{k^3 ts_i c C_c}{104} = \$9303$$

Now, this is a very large value for K , considering the cost of the item. Of course, it is possible that such a value of K is correct but it is certainly extremely unlikely. This is the reason we commented on the probable absurdity of the determined value of k . Now that we have derived the expression for the optimal value of k we can use it directly if we know K . Thus, suppose that for our item, K is determined to be \$1000, still a large value. Then a direct application of our equation for k gives:

$$k = \sqrt[3]{\frac{(104)(1000)}{(5.89)(8.66)(5)(.12)}} = 15.0$$

with a resulting reserve stock of 129.9.

For any one of the inequalities which is appropriate a corresponding analysis can be made. If we know both the second and the fourth moments about the mean of our demand distribution—summarized by a knowledge of γ_2 —then a repetition of the previous argument produces the following equation

$$\frac{4k(k^2 - 1)}{[(k^2 - 1)^2 + \gamma_{2,1} + 2]^2} = \frac{ts_i c C_c}{52K(\gamma_{2,1} + 2)}$$

This equation cannot be conveniently solved explicitly for k but, for any specific values for the other quantities, it can be solved and will give the optimal k for the inventory problem in question. For example, if we know that $\gamma_2 = 0$ for our item we calculate that k is very nearly 6.3 or a reserve stock of about 54.6.

Let us now consider the exact solution for this case. We will use the total cost equation based on the Tchebycheff inequality as an example. For the exact solution we must write one total cost equation which expresses both ordering costs and reserve stock costs. We will write this equation in terms of t , the number of weeks covered by one order:

$$T.C_t \leq \frac{52C_r}{t} + \frac{t\bar{z}cC_c}{2} + ks_1cC_c + \frac{52K}{t} \frac{1}{k^2}$$

with derivatives giving, with respect to t and k

$$t = \frac{104K}{k^3s_1cC_c} \quad \text{and} \quad t^2 = \frac{104(C_rk^2 + K)}{k^2\bar{z}cC_c}$$

Thus, we can equate these values for t and solve for k :

$$\frac{104\bar{z}cC_cK^2}{(s_1cC_c)^2} = k^4(C_rk^2 + K)$$

Once again, this is not an equation suitable for an explicit expression for k but the equation can be solved for any specified set of values for the other quantities. Thus, for our example, using $K = \$1000$, we calculate that $k = 14.04$. This gives $t = 7.23$. Thus, the optimal policy calls for an order size of $7.23(50) = 361.5$ units and a reserve stock of 121.6 units. As compared to the approximate analysis there has been a relatively small change, since for the approximate analysis the order period was found to be 5.89 and the reserve stock was 129.9. The approximate analysis resulted in a policy which guaranteed that the total cost per year would never exceed \$293.82 while the exact analysis produces a maximum total cost per year of \$289.91.

The other inequalities can be utilized in a similar way to determine exact analyses of this kind of situation. The algebra involved is straightforward but the numerical calculations can become onerous. However, it should be remembered that it is relatively easy to prepare ancillary tables which permit the rapid calculation of the optimal policy for any one of the inequalities which is most frequently appropriate in a specific organization. Furthermore, if many items are involved the numerical calculations would be handled by a computer in a relatively straightforward manner.

The analyses needed for the third case, where there is a per unit out-of-stock cost, are very similar in structure to those for the second case. The major difference is that we must use the somewhat more difficult methods developed in Section 25. This makes the work more arduous but does not affect in any way the basic conceptual framework of the analysis.

Using the approximate method of analysis and using the Tchebycheff

inequality as an example, we have the total cost equation—based on the results of Section 25—as follows:

$$T.C_r \leq k s_i c C_c + \frac{52 C_u s_i}{t} \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

This equation must be minimized with regard to k . We find the following equation:

$$\frac{t c C_c}{26 C_u} k^4 = 2k^2 + 2k + 1$$

The positive root of this equation is the desired optimal value of k . Assuming that for our item $C_u = \$50$, we find that k is very nearly 27.5 or a reserve stock of 238.2 units. This amount of reserve stock would result in a total cost less than or equal to \$461.60 per year.

For the exact analysis of the same example we would start with the total cost equation:

$$T.C_i \leq \frac{52 C_r}{t} + \frac{t z c C_c}{2} + k s_i c C_c + \frac{52 C_u s_i}{t} \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

This will be a convenient place to indicate a simplification that can sometimes be made in this kind of equation. We observe that k was quite large in the approximate solution so it appears legitimate to assume that it will be fairly large in the true optimal solution. If this is so the last term in the parentheses of the total cost equation becomes insignificant. Thus, for $k = 27.5$ we have $1/6k^3$ is less than .00001. Under these circumstances a perfectly satisfactory approximation can be achieved by dropping the $1/6k^3$ term from the equation. So doing, taking the two derivatives with respect to t and to k , and performing some algebraic manipulations gives the following equation:

$$\frac{26 C_u^2 z}{c C_c} \left(\frac{1}{k^2} + \frac{1}{k^3} \right)^2 = C_r + C_u s_i \left(\frac{1}{k} + \frac{1}{k^2} \right)$$

This is another awkward equation which needs to be solved for the specific problem. In our case, and using $C_u = \$50$ we find $k = 20.9$ or a reserve stock of 181.0. For the solution to the above equation, t is calculated from

$$t = \frac{52 C_u}{c C_c} \left(\frac{1}{k^2} + \frac{1}{k^3} \right)$$

For our example, $t = 10.4$. Thus, this policy would require average orders of $x = 520$. The total cost for this policy would be less than or equal to \$420.68.

These, then, are the methods which can be used when partial information is available and it is desired to determine a Q -system. We have only given one example of each of the various approaches but in each case the procedure would be exactly the same if another inequality were used than the one we chose for our example. As a general rule, the numerical computations are annoyingly lengthy but this is only because we used straightforward numerical

methods in each case. Any of these equations can be solved with completely adequate accuracy and quickly if nomograms or ancillary tables are prepared. Obviously, this is not worth doing unless there are a number of items for which the same kinds of calculations must be made. But, in this event, the procedures are quite rapid and efficacious.

46. P-SYSTEMS FOR INVENTORY CONTROL

It will be remembered that the *P*-system is characterized by having a fixed order period and a variable order quantity. As was discussed previously, the determination of the optimal reserve stock for the *P*-system should, theoretically, require consideration of the complex interactions among successive orders. However, here we will continue to make the generally accepted simplifying assumption that the reserve stock can be satisfactorily approximated by basing it on the demand fluctuations for the order period plus the lead time. The analysis of this system by distribution-free methods is quite parallel to the analyses of the *Q*-system in the preceding section. Therefore, we will not give so many examples.

We can begin by considering the distribution-free analogy to the approximate method of calculation. The approximation consists of using the optimal lot-size equation to calculate the order size and, hence, order period. Then the reserve stock for the order period plus lead time is determined separately. For the item which we are using as an example the order period plus lead time is 8.89 weeks. At this point we will offer a proviso which will apply throughout this section. In strict fact, convolutions are only defined for integral multiples of the basic period over which the demand distribution is known. Therefore, one should round the given order period off to the nearest integral value—or the least expensive, by the methods we discussed in Chapter 4—and use this for the subsequent calculations. However, we prefer to use the actual nonintegral order period instead. For the approximate method it would be just as easy to use the rounded off value and those who prefer may do so. But for the exact analyses it is necessary to use the exact, non-rounded value to avoid the mathematical horrors which otherwise result. The use of the exact value in these cases can be considered to be an approximation to the use of the rounded values.

Returning to our example, we may note that in the preceding chapter when we discussed *P*-systems we considered alternative ways of calculating the cost of the reserve stock. In our distribution-free analyses we will only use the simpler method—assuming that the reserve stock is carried all year round. Alternative ways of calculating the carrying costs for the reserve stock can be introduced but at the cost of additional arithmetical complexity. Using the Tchebycheff inequality we can immediately write for the reserve stock cost equation:

$$T.C_r \leq ks_r C_c + \frac{52K}{tk^2}$$

In this equation $s_i = s_{8.89} = 5\sqrt{8.89} = 14.91$. This equation, which is based on a fixed out-of-stock cost, K , is precisely the same as the analogical one in the preceding section. The difference, of course, resides in the longer period for which the reserve stock must be protection against demand fluctuations. The minimizing expression for k is the same:

$$k = \sqrt[3]{\frac{104K}{ts_i c C_c}}$$

For our example we calculate, assuming $K = \$1000$, that $k = 10.9$. The reserve stock is therefore 163.0. Total costs for this policy are \$324.36. Comparing this with the analogous maximum total cost for the Q -system when the approximation was used, we find that the latter was \$293.82. This is in accord with our expectation that the Q -system will have lower total costs than the P -system.

It is not necessary to give an example of the exact method of analysis for this case since we show in the Appendix that the exact analysis gives the same expressions for t and k which we have obtained by the approximate analysis. To the degree that coincidences can occur in mathematical analysis we can say that this is a coincidence but it does spare us the necessity of having to consider the exact analysis for this case.

The approximate method for per unit out-of-stock costs proceeds exactly as for the corresponding analysis of the Q -system. Basing our analysis on the results of Section 25, we can write the total cost equation for the reserve stock as

$$T.C_r \leq ks_i c C_c + \frac{52C_u s_i}{t} \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

the same equation as for the Q -system analysis. The difference, of course, is in the value of s_i . The equation for the minimizing value of k is also the same

$$\frac{tcC_c}{26C_u} k^4 = 2k^2 + 2k + 1$$

This equation is exactly the same in appearance as the corresponding equation for the Q -system. The difference, of course, is that here t includes the lead time and, hence, equals 8.89 instead of 5.89, as in the Q -system analysis. With $C_u = \$50$, the solution of this equation is $k = 22.6$. This gives a reserve stock of 337.1. The total cost of this policy will be less than or equal to \$575.94.

For the exact analysis of the same example we would start with the total cost equation:

$$T.C. = \frac{52zC_r}{x} + \frac{xcC_c}{2} + kscC_c \sqrt{\frac{x}{z}} + LT + \frac{52C_u s}{\sqrt{\frac{x}{z}} + LT} \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

Of all of the equations we have written, this one is the most frightful and monstrous. It is ugly, unpleasant, and unesthetic. However, it will give the

optimal inventory policy under the stated conditions. Actually, it is less difficult to minimize than many of the others we have used. Probably the simplest way is to take the derivative with respect to k and solve for x , giving

$$x = z \left[\frac{52C_u}{cC_c} \left(\frac{1}{k^2} + \frac{1}{k^3} + \frac{1}{2k^4} \right) - LT \right]$$

Now we can try various values of k , finding the corresponding values of x , and work directly with the total cost equation till we find the minimum to a sufficient level of approximation. Proceeding in this fashion we discover that the total costs for k between 22 and 23 are all less than \$577. On both sides of this range of values for k the total costs go up. Therefore, the optimal value of k is in this range and it makes little difference exactly where. As a result we can assume that $k = 22.5$, with the same solution we found when we used the approximation.

These examples should suffice to indicate how partial information can be utilized in determining optimal P -systems. There can be no doubt but that some of the other inequalities which may be used can lead, for specific models, to awkward equations—as, indeed, did the Tchebycheff inequality in one or two cases above. However, we will emphasize for the last time that this is a very minor matter. Nomograms and similar devices can always be constructed to help in quickly and accurately determining the solutions to the equations which may be derived. The essential point is that it is possible to utilize partial information in constructing inventory control systems which are optimal within the bounds of the given information and within the given class of systems.

47. THE USE OF CENTRAL WAREHOUSES

The analyses of the two preceding sections are relatively straightforward extensions of the approaches we developed in the chapter dealing with static inventory problems under uncertainty. It will be worthwhile now to illustrate how distribution-free analyses can be utilized in the development of solutions to other kinds of inventory problems. The example we will use will be the problem of the utilization of a central warehouse with regard to reserve stocks. Suppose there are n warehouses and a central warehouse. Some items will typically be stocked by all of the warehouses—and, of course, we can always redefine n if not all warehouses stock an item which we are interested in analyzing. Each warehouse will have to maintain some reserve stock for the item in question. The possibility exists that each warehouse might assign some portion of its reserve stock to the central warehouse and that the central warehouse could maintain the same total level of protection while actually carrying less stock than the total of the amounts assigned to it by the individual warehouses. This would result in a saving of the given amount of carrying charges. Balanced against this would be any additional costs for extra handling and extra transportation. However, it may be that a net saving could result. This is the situation which we want to investigate.

The possibility of such a saving results from the fact that the standard deviation of a sum of random variables goes up as the square root of the number of variables in the sum. This means that a given level of protection against demand fluctuations is relatively cheaper for a sum of such variables than it is if the same level of protection is maintained for each variable separately. We have already seen examples of this in terms of the convolutions of demand distributions, and the case of individual warehouses is exactly the same.

We will suppose, then, for simplicity, that the demand at each warehouse for the given item is the same, with mean equal to $\bar{x}$ and standard deviation equal to s , both on a monthly basis. We might use any one of our inequalities as the basis for our analysis, depending on the nature of our partial information. As a typical example we will assume that we know that the demand distribution—the same for each warehouse—is unimodal and symmetrical. Finally, we will assume that there is a fixed out-of-stock cost, K . The warehouses use a P -system of inventory control and the policies are determined by the approximate method. In other words, the order quantity is determined from

$$x_0 = \sqrt{\frac{24\bar{x}C_r}{cC_c}}$$

and there are $n = 12\bar{x}/x_0$ orders per year. If the lead time is LT months, it follows that the time period on which the reserve stock will be based will be $t = \frac{x_0}{\bar{x}} + LT$. For each warehouse this means that the standard deviation of demand for the order period will be given by

$$s_t = s\sqrt{t}$$

For the stated information about the demand distribution we would use the inequality

$$P[(y - \bar{x}) \geq ks_t] \leq \frac{2}{9k^2}$$

However, it will be more convenient for us to express this directly in terms of the reserve stock, r . The same inequality in terms of r is:

$$P[(y - \bar{x}) \geq r] \leq \frac{2s_t^2}{9r^2}$$

The maximum probability of going of-stock is, of course, the right-hand side of this expression. We will designate this probability as α . Therefore,

$$\alpha = \frac{2s_t^2}{9r^2}$$

The reserve stock would be determined by minimizing

$$T.C_r = rcC_c + \left(\frac{12K}{t}\right) \left(\frac{2s_i^2}{9r^2}\right)$$

Substituting $s_i = s\sqrt{t}$ this becomes

$$T.C_r = rcC_c + \frac{8Ks^2}{3r^2}$$

This is minimized by

$$r = \sqrt[3]{\frac{16Ks^2}{3cC_c}}$$

which imputes a value to α of

$$\alpha = \left(\frac{2}{3}\right)^{1/3} \left(\frac{t}{12}\right) \left(\frac{cC_c s}{K}\right)^{2/3}$$

This, then, determines the reserve stock policy which each warehouse will follow if left to its own devices. We want now to investigate whether any savings are possible of the sort we discussed in the first paragraph of this section. Obviously, there cannot be any savings if there is only one warehouse apart from the central warehouse. We will assume that every warehouse will carry sufficient stock to meet its own average demand because it is clear that no savings could result from having the central warehouse carry part of this stock. Our analysis will be based on the plausible assumption that no warehouse will be entitled to withdraw from the central warehouse more reserve stock than it assigned to the central warehouse.

It is obvious that if a warehouse assigns part of the reserve stock determined above to the central warehouse then the probability of an out-of-stock will increase for the assigning warehouse. Whereas the probability of an out-of-stock was originally α it will now be some probability greater than this, α_1 . If the warehouse goes out-of-stock it will call back its assigned stock from the central warehouse. The probability that it will have to do this will be designated β . Clearly,

$$\beta = \alpha_1 - \alpha$$

It will generally be the case that β is small. This being so, the probability $\phi(i)$ that i warehouses will, in the same order period, call back stock from the central warehouse will be given by the Poisson distribution:

$$\phi(i) = \frac{(n\beta)^i e^{-n\beta}}{i!} \quad \text{for } i = 0, 1, \dots, n$$

where there are n warehouses.

We want to find some amount which can be assigned by the individual warehouses to the central warehouse such that the same level of protection can be provided for the group of warehouses but with savings in carrying

costs. Each warehouse will send an amount of reserve stock corresponding to β . The central warehouse must maintain a total probability of being out-of-stock of α . If the sum of the probabilities, $\phi(i)$, from j to n is less than or equal to α the central warehouse will have maintained the necessary level of protection and will simultaneously have effected a savings of $(n - j)$ times the average number of units sent by each warehouse. Thus, we want

$$\sum_{i=j}^{i=n} \phi(i) = \sum_{i=j}^{i=n} \frac{(n\beta)^i e^{-n\beta}}{i!} \leq \alpha$$

This equation will ordinarily be satisfied for various combinations of values of j and β . In general, β will decrease as $(n - j)$ decreases. We want to find that combination of β and $(n - j)$ which maximizes the savings.

We need to know the number of units to be assigned to the central warehouse which corresponds to a particular value of β . From

$$\alpha = \frac{2s_i^2}{9r^2}$$

we find

$$r = \sqrt{\frac{2s_i^2}{9\alpha}} = .471s_i \sqrt{\frac{1}{\alpha}}$$

Exactly as r depends on α and represents the total reserve stock which the individual warehouse would carry if it carried all of its own reserve, so there is some r_1 corresponding to α_1 and representing the reserve stock the individual warehouse will carry if the central warehouse carries part of its reserve. Obviously, we must have $r_1 < r$. Since $\alpha_1 = \alpha + \beta$ we have

$$r_1 = .471s_i \sqrt{\frac{1}{\alpha + \beta}}$$

The amount of stock, in units, corresponding to β is, then,

$$r - r_1 = .471s_i \left(\sqrt{\frac{1}{\alpha}} - \sqrt{\frac{1}{\alpha + \beta}} \right)$$

The savings in units in reserve stock will, therefore, be

$$(n - j)(r - r_1) = (n - j)(.471s_i) \left(\sqrt{\frac{1}{\alpha}} - \sqrt{\frac{1}{\alpha + \beta}} \right)$$

It is this quantity we want to maximize by a choice of $(n - j)$ and β . In order to do this we need tables of the cumulative Poisson distribution and a little calculation.

Let us illustrate the procedure with an example. Suppose we have the following data:

$c = \$16000$	$\bar{x} = 2$ per month
$C_c = .12$	$s = 1$
$C_r = \$50$	$n = 10$ warehouses
$K = \$50000$	$LT = 3$ months

We calculate that $x = 11.2$, $t = 8.6$ months, $r = 5.18$, and $\alpha = .0722$. This means that we want to consider those Poisson distributions for which

$$\sum_{i=j}^n \phi(i) \leq .0722$$

Let us begin by putting $j = 1$. For what Poisson distribution is the probability of from 1 to 10 successes, inclusive, $= .0722$. From the tables the answer to this is that the Poisson distribution with mean of .075 meets this requirement. Thus, $n\beta = .075$ and $\beta = .0075$. Therefore, $\alpha + \beta = .0797$ and $\sqrt{\frac{1}{\alpha}} - \sqrt{\frac{1}{\alpha + \beta}} = .18$. The total savings in units for this case would therefore be $(n - j)(.471)(2.93)(.18) = 2.24$ units. Continuing in this way we find that the maximum savings occur for $j = 4$. This gives $\beta = .155$ and a total savings of 13.41 units.

The expected number of withdrawals per order period would be $n\beta = 1.55$. The amount assigned by each warehouse to the central warehouse would be

$$(.471s_1) \left(\sqrt{\frac{1}{\alpha}} - \sqrt{\frac{1}{\alpha + \beta}} \right) = 2.24$$

The central warehouse can maintain the given level of protection with only 9.0 units in stock instead of the 22.4 required if each warehouse maintained all of its own reserve. This represents a saving of about \$25,700 per year in carrying costs.

The opposing costs consist of the requisition costs for withdrawals and extra handling and transportation costs. Since we know the average number of withdrawals per period is 1.55, or 2.16 per year, it is possible to determine these opposing costs and see to what degree they subtract from the savings. Obviously, if they are greater than the savings this policy would not be put into effect.

PROBLEMS

1. An item has average demand of 100 units per week and standard deviation of demand equal to 10. Order cost is \$10 and carrying cost is 12% per year. The price of the item is \$6. Assume a fixed understock cost of \$100. The lead time is three weeks.
 - (a) Use the Tchebycheff inequality and the approximate method of analysis to determine the optimal Q -system for the item.
 - (b) Assume that the demand distribution is unimodal and symmetrical. Use the approximate analysis to determine the optimal Q -system.
 - (c) Assume the demand distribution is unimodal with a mode at 90 units. Use the approximate analysis to determine the optimal Q -system.
 - (d) Suppose that in each of the three preceding parts the understock cost was not given. Instead, the allowable percentage of understocks was set at .001. What understock cost is imputed in each case?

- (e) Suppose it is known that $\gamma_2 = 0$. Use the approximate method of analysis to determine the optimal Q -system.
 - (f) Under the assumptions of (a), use the exact method of analysis to find the optimal Q -system.
 - (g) Assume a per unit understock cost of \$10. Use the Tchebycheff inequality and the approximate method of analysis to determine the optimal Q -system.
2. Determine the optimal P -system for the item of Problem 1. Parts (a)–(g) of this problem correspond to the same parts of Problem 1.

- (b) Suppose it is known that $x \in U$. The corresponding method of analysis to determine the optimal Q -system.
- (c) Under the assumptions of (a) and the exact method of analysis to find the optimal Q -system.
- (d) Assume a per unit overhead cost of 100. Is the Total Cost inequality and the approximate method of analysis to determine the optimal Q -system.
2. Determine the optimal Q -system for the case of Problem 1. Point out if of this problem corresponds to the same parts of Problem 1.

part two

implementation phase: the theory of practice

the evaluation study: design of the inventory study

A properly chosen inventory system will minimize costs by balancing the antagonistic cost systems. If an important cost or set of such costs is overlooked or purposely ignored, the analysis produces incorrect conclusions. Many times the evidence of such errors of omission or judgement is lost—hopelessly entangled and absorbed in the amalgamated overhead of non-specific costs which are felt but not recognized by the company.

In particular, the set of costs which we have named systemic costs are susceptible to formal analysis but they are likely to be excluded from even the most comprehensive inventory analyses. We must ask why this is so. Is it the fact that they don't really matter? Perhaps the measurement of systemic costs presents insuperable problems? Or maybe, management prefers to use its intuition with respect to such decisions because of habit and tradition? Let us for the moment reserve judgement on the matter.

First of all, we shall explain the structure of Part 2 of this book. It consists of three chapters which treat in turn:

- (1) The evaluation study: design of the inventory study;
- (2) The inventory study: design of decision procedures;
- (3) The data-processing study: design of operating procedures.

We recognize that these three phases coexist in any inventory study. The boundaries are arbitrary and since no well-defined dividing lines exist it is sometimes reasonable to categorize a topic as belonging with equal right to more than one of our divisions. But this fact does not detract from the usefulness of the structure. It has the advantage of centering attention on relatively homogeneous sets of systemic costs. It is also intended to do away with the prevalent notion that the subject of inventory is totally defined by listing all possible inventory models. Certainly, models are at the very heart

of our inventory study, but without recognition of the phases of the study and the way in which theory interacts with practice, the models are like giants with feet of clay. That is why Part 2 is called the implementation phase. It is concerned with both theory and practice. In the process of studying implementation we discover that theory relating to subjects other than inventory models concerns us as well as the theory of inventory itself. We find that dependencies exist which bind the consequences of decisions in one area to the results obtained in another. Having shed some light on our direction and approach, let us proceed to the first point—the evaluation study.

Inventory studies cost money and they have a definite utility. Unfortunately, both the cost of the study and the utility returned are most easily determinable *a posteriori*. It is our purpose in this chapter to find out what can be done about obtaining *a priori* estimates of these opposing factors. Based on such estimates a decision will be made to study—or not to study—the system. Either decision can be wrong. That is obvious. What is not obvious is that both decisions can be wrong. It follows then that a dichotomy (to study or not to study) is not sufficient. A large range of study possibilities exists just as a large range of inventory models, forecasting methods, and data-processing procedures exists. We will direct our attention in this chapter to the various controls which management can exercise through study design over both the costs of studies and the benefits derived.

48. ELEMENTS OF THE INVENTORY STUDY

What is entailed in the study of an inventory system? To begin with, a decision is required. This is a decision spawned by dissatisfaction with the present state of affairs and based on the belief that a substantial savings can be obtained by allocating remedial efforts to the inventory area. It is certainly apparent that management would turn its efforts in other directions if the likelihood of a higher payoff existed in other areas. Investigations to improve efficiency and effectiveness are limited by the resources of time, adequate personnel, and money which management can call upon. Furthermore, careful planning and supervision are required to insure that an investigation will not create operating and personnel problems. The care entailed to avoid disruption has a definite associated cost which in its own right should be balanced against the costs of disruptions. Such costs are indeed tenuous. They are not available from accounting records. One can readily understand why management ignores formal analysis and prefers to use its intuition and experience in determining which areas of the company's operations will be most fruitful for investigation.

Unless management has definite reason for doing otherwise, it is advisable that an evaluation of the present system should precede the design of a new system. In this way it can be determined whether or not the present system should be changed—in what way it should be changed—and to what extent it should be changed.

The cost of the evaluation study is nonrecoverable if at its conclusion the

decision is made to stop. This decision to stop is always logical if the total cost savings of a new system will not provide a better return on the capital invested in the study than on that capital invested in another way. Thus, Criterion 1:

$$\frac{\text{Annual savings obtained from the inventory study}}{\text{Cost of the inventory study}} \geq C_c$$

where C_c is equal to the best return on investment that management can get if it invests its money in another way. It is the purpose of the evaluation study to provide *a priori* estimates of both of the factors in the left-hand side of the equation above. The estimation of C_c is management's responsibility. It is, however, within the scope of management science to help the executive determine an appropriate C_c . It should be noted that this is the same C_c which appeared in the extensive treatment of inventory models in the first part of this book; i.e., an interest rate per unit of time. The savings and costs are computed to cover the same interval.

The only way that the cost of an inventory study is recoverable is through the time stream of savings. Therefore, the criterion which is shown above may not be at all satisfactory if the savings derived from the study are short-lived. Such a condition might prevail if the results of the inventory study are not transferable to new product lines which are intended to replace the present product mix. Under such circumstances, the company may prefer a different criterion. For example, a three year breakeven period may be the objective. We can then write Criterion 2:

$$\frac{\text{Required savings at present worth (for } n \text{ years)}}{\text{Cost of the inventory study}} \geq 1$$

where $n = 3$.

This criterion can also be used when the benefits of an inventory study are applicable to the company over a long period of time. Generally the fixed costs of an inventory study which go primarily to support an investigating group, do not have to be repeated. The group determines and sets up methods of analysis and rules for decision-making which work for a variety of cases—all of which are foreseeable in the company's future. Variable costs of the study are likely to be repeated as new items join the product line and as presently carried items are updated.* This fact can be taken into account by computing the cost of the study as a fixed cost plus the present worth (for n years) of the variable components.

It is also possible to use a very long-range criterion. Here the management states that they will be prepared to accept an inventory study if the present worth (for lifetime) savings are as follows, Criterion 3:

$$\frac{\text{Required savings at present worth (for lifetime)}}{\text{Cost of inventory study}} \geq f$$

where $f > 1$. Here too, we can treat the study cost as being composed of

* The subject of updating is treated in Chapter 9.

fixed and variable components. That is, a fixed cost plus the present worth of the variable components based on a lifetime stream.

It is obvious that each criterion we have given could lead to different decisions.* Our purpose here has been to firmly establish that the inventory study starts in the anonymous headwater of management prerogative. And, in the exacting tradition of the Order of the Garter, "Honi soit qui mal y pense." But on the other hand, it is management's duty to supply the criteria necessary to render their objectives unambiguous. For comparison of the different criteria we can make use of the following relations:

$$\text{Criterion 1:} \quad \frac{\psi}{\chi_1 + \chi_2} \geq C_c$$

$$\text{Criterion 2:} \quad \frac{\rho_n \psi}{\chi_1 + \rho_n \chi_2} \geq 1$$

$$\text{Criterion 3:} \quad \frac{\mu \psi}{\chi_1 + \mu \chi_2} \geq f$$

ψ = savings per year,

χ_1 = fixed study costs,

χ_2 = variable study costs per year,

ρ_n = present worth of \$1 obtained at the end of each year for n years,

$$= \left[\frac{(1 + C_c)^n - 1}{C_c(1 + C_c)^n} \right]$$

μ = present worth of \$1 based on a lifetime stream,

$$= \left(\frac{1 + C_c}{C_c} \right)$$

C_c = carrying rate per year.

Once the decision has been made to enter the evaluation phase of the inventory study and management has specified its criterion, the remaining major elements of the inventory study are as follows:

(1) A representative sample of items is analyzed with respect to costs, demand, lead time, etc.

(2) Based on the simplest models, theoretical performance is compared with the actual performance. Expected savings and study cost are derived and evaluated by management's stated criterion.

(3) The decision is made to continue or to stop the study.

(4) A complete analysis, subject to grouping rules (see Section 51), is made for all items.

(5) A minimum but sufficient number of inventory models are designed to cover the various company activities.

(6) The proposed systems are tested. Indicated changes are made until satisfactory performance is obtained.

* See Problem 1 in the problem section of this chapter.

These steps are, of course, only representative. It is not unusual for an inventory study to proceed sequentially in both breadth and depth. By breadth, we mean that only a portion of the total number of items are placed under the new controls and gradually the remainder are added. By depth we mean that a simple set of inventory rules is introduced at first. These are gradually replaced by more complex rules.

49. INTUITIVE AND JUDGEMENTAL APPROACH

The cheapest method of evaluation is by means of executive judgement. In a company which possesses an effective communication system, a well-designed record system, and an alert management, it is quite possible that reasonable estimates can be made. Management's experience with understocking, overstocking, spoilage, discounts, load on the ordering department, characteristic lead times, and the effects of their variability on inventory levels, must be complete. At the same time, management must be familiar enough with inventory theory to be able to estimate the improvement that would be possible if different inventory policies were adopted. Based on the experience of other companies, it would be possible to estimate how much real improvement might be obtained by altering the present system. Lastly, a reasonable figure could be assigned to the cost of such a study.

Usually, management cannot satisfactorily perform the functions above. Their attempt to do so has a close affinity to the use of a Ouija board. It should be noted, however, that while large* companies with substantial investments in inventory can ill afford the services of a Ouija-board approach, smaller* companies are even more affected by errors in estimation. Consider a company with a million dollar inventory. Assume that they estimate a 5% savings per year resulting from the inventory study which it is estimated will cost \$90,000. Then,

$$\frac{\text{Savings}}{\text{Cost}} = \frac{50,000}{90,000} = 55\frac{5}{9} \%$$

It is unlikely that any management would ignore this opportunity. Let us now compare the result above with that of a company which maintains an inventory of \$100,000. We will presume that the cost of the study will be only \$30,000. (Note, economies of scale exist which prevent us from estimating the cost of the study at \$90,000.) Then,

$$\frac{\text{Savings}}{\text{Cost}} = \frac{5000}{30,000} = 16\frac{2}{3} \%$$

Thus, although both the large and small companies can effect the same percent savings from an inventory study, it represents much less freed capital for the smaller company and they must be more precise in reaching their estimates.

* By large and small in this context we refer to the size of the inventory investment.

Another factor which works against the smaller* company is that the probable level of savings which can be obtained from an inventory study will be lower than that which can be expected by the larger* company. This results from the exponential increase in complexity which occurs with increasing organization size. The executives in a smaller company are closer to the total system. In this sense they have relatively more information and can exercise greater control. Thus, through a circumstantial, structural advantage the smaller company will have less waste and inefficiency which can be removed by a formal inventory system. We, of course, are presuming that this comparison deals with companies which possess comparable levels of executive abilities.

Since the savings percent tends to be smaller and is applied to a smaller base of investment and since the cost of the study has a decelerating increase as the size of the inventory investment increases, it behooves the smaller companies to exercise greater caution in fixing their decision estimates. Consequently, one might expect to find that the executives of smaller companies have less sanguine attitudes with respect to the extent of untold benefits to be derived from inventory studies. They are also more likely to use intuition and judgement alone in reaching a decision about undertaking a study. The soundness of this position can only be contested by a rational procedure for evaluating the benefits likely to be obtained from an inventory study.

50. SIZE OF INVENTORY INVESTMENT AND NUMBER OF ITEMS CARRIED

Reference to the size of a company is too ambiguous to be a genuinely useful classification for companies with respect to their inventory problems. At the least, we should specify each company's position in terms of the size of their inventory investment and the number of different items which are involved. A company which deals with just one item can afford to obtain a great deal of information about the demand pattern and lead time for that item. It is also likely that executive judgement will be very efficient in determining the best possible order and storage procedures for the item without having recourse to formal studies. However, if the inventory investment in the item is very large, then small errors in judgement will be magnified to produce significant opportunity costs. For example, let

$$x_o = \sqrt{\frac{2zC_r}{cC_c}}$$

be the "true" optimal order quantity; i.e., the estimates of the parameters are correct. Then,

$$T.C_{opt} = \sqrt{2zcC_rC_c}$$

* By larger and smaller in this context we refer to the size of the inventory organization. It is also logical that complexity tends to increase as the number of items, locations demanding items, and locations storing items increase.

Now assume an error in the estimation of C_c . Let the estimate be kC_c . Then, we would derive an order quantity,

$$x_1 = \sqrt{\frac{2zC_r}{ckC_c}}$$

which would yield a total cost,

$$T.C_1 = \left(\sqrt{\frac{2zC_r}{ckC_c}} \right) \left(\frac{cC_c}{2} \right) + \left(\sqrt{\frac{ckC_c}{2zC_r}} \right) (zC_r)$$

or
$$T.C_1 = \sqrt{\frac{zcC_cC_r}{2k}} + \sqrt{\frac{zckC_rC_c}{2}} = \frac{1}{2} \left(1 + \frac{1}{k} \right) k^{1/2} (T.C_{opt})$$

and

$$\frac{T.C_1}{T.C_{opt}} = \frac{k+1}{2\sqrt{k}}$$

also, the potential savings per period would be:

$$T.C_1 - T.C_{opt} = T.C_1 \left(1 - \frac{2\sqrt{k}}{k+1} \right)$$

Figure 7-1 illustrates the way in which the potential percent savings per period of time, ϕ_k , changes as the error, k , varies in size.*

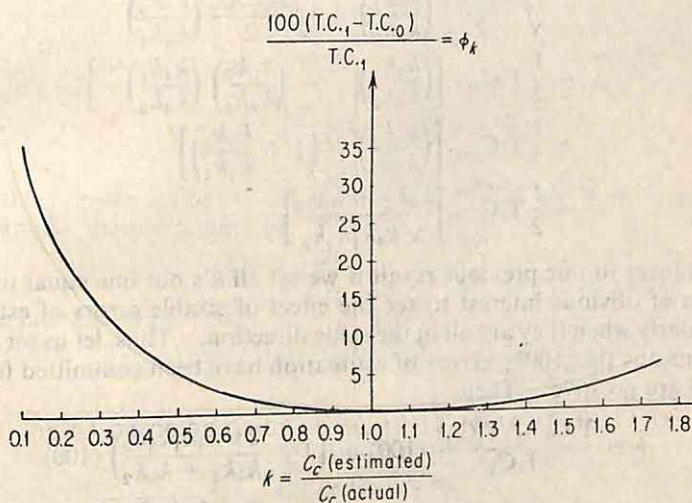


FIG. 7-1 Potential percent savings vs. error ratio in estimating carrying cost rate, C_c .

* We have previously performed a similar sensitivity analysis utilizing two errors of estimation (see pages 83-84). It will be observed that the results are comparable—as they should be—when k_1 , in the previous analysis is set equal to one.

We now see graphically that the potential saving which can be obtained from the study of costs is highly insensitive* to the size of the errors which exist in the estimation of the cost factors. However, the base size of the T.C. varies directly with the square root of zc . Thus, in addition to high demand items with high prices, both low demand items with high prices and high demand items with low prices can produce significantly high values of T.C. When even small percent savings are applied to a large value of T.C., the dollar savings can be substantial. Consequently, it is less meaningful to talk about large and small companies and more appropriate to classify companies by their total zc (called $T.zc$) and their characteristic level of T.C.

We can easily extend our analysis of sensitivity to errors of estimation to any other parameter which must normally be estimated in the course of an inventory study. To be entirely general, we will presume that estimation errors can and do occur in C_c , C_r , c , and z . As before, we will let the un- k 'd factors be the "true" values and the estimates will be k_1C_r , k_2C_c , k_3c , and k_4z . Then,

$$\begin{aligned}
 x_1 &= \sqrt{\frac{2(k_4z)(k_1C_r)}{(k_3c)(k_2C_c)}} \\
 \text{and } T.C_1 &= \sqrt{\frac{2(k_4z)(k_1C_r)}{(k_3c)(k_2C_c)}} \left(\frac{cC_c}{2} \right) + (zC_r) \sqrt{\frac{(k_3c)(k_2C_c)}{2(k_4z)(k_1C_r)}} \\
 &= \sqrt{\frac{zcC_rC_c}{2} \left(\frac{k_4k_1}{k_3k_2} \right)} + \sqrt{\frac{zcC_rC_c}{2} \left(\frac{k_3k_2}{k_4k_1} \right)} \\
 &= \frac{1}{2} T.C_{\text{opt}} \left[\left(\frac{k_4k_1}{k_3k_2} \right)^{1/2} + \left(\frac{k_3k_2}{k_4k_1} \right)^{1/2} \right] \\
 &= \frac{1}{2} T.C_{\text{opt}} \left[\left(\frac{k_4k_1}{k_3k_2} \right)^{1/2} \left(1 + \frac{k_3k_2}{k_4k_1} \right) \right] \\
 &= \frac{1}{2} T.C_{\text{opt}} \left[\frac{k_4k_1 + k_3k_2}{\sqrt{k_4k_1k_3k_2}} \right]
 \end{aligned}$$

This reduces to our previous result if we set all k 's but one equal to unity.

It is of obvious interest to see the effect of sizable errors of estimation—particularly when they are all in the same direction. Thus, let us set all $k = 2$ which means that 100% errors of estimation have been committed four times and all are positive. Then,

$$\begin{aligned}
 \phi &= \frac{T.C_1 - T.C_{\text{opt}}}{T.C_1} (100) = \left(1 - \frac{2\sqrt{k_4k_1k_3k_2}}{k_4k_1 + k_3k_2} \right) (100) \\
 &= \left[1 - \frac{2\sqrt{(2)(2)(2)(2)}}{(2)(2) + (2)(2)} \right] (100) \\
 &= 0
 \end{aligned}$$

* The sensitivity is greater for underestimation than overestimation, but within a reasonable range of underestimation it is fair to say the response lacks sensitivity. It should also be pointed out that the likelihood of severe underestimation is much less than of severe overestimation.

That is, $T.C_1 = T.C_{opt}$ in spite of the enormity of the errors of estimation. These errors have completely cancelled each other out.

Upon reflection, it is obvious that this is the only logical outcome. The opposing costs have merely balanced each other, which would occur for all equal multiples of the factors. What happens then when the term $k_4 k_1$ is different from $k_3 k_2$? Let us try the following values: $k_1 = k_4 = \frac{1}{2}$ and $k_2 = k_3 = 2$.

$$\phi = \left[1 - \frac{2\sqrt{(.25)(4)}}{4.25} \right] (100) = 53\%$$

Here we have two errors of 100% increase and two errors whose magnitudes are 50% in the direction of decrease. The result is that 53% savings in T.C. are obtainable by means of an inventory study which uncovers and rectifies the estimation errors.

But we must note that errors in estimating demand and ordering costs are of the same type. Similarly, errors in estimating price and carrying cost rates are of the same type. As the differences between these two types of errors grow, the percent savings available increases.

To round out our present discussion on potential savings that can result from an inventory study, let us assume that each of the four k factors are independent and distributed in the following fashion.

k	$P(k)$
1/4	1/16
1/2	3/16
1	1/2
3/2	3/16
2	1/16

From this we can compute the distribution of potential percent savings. For example, the probability of

$$(k = \frac{1}{4})(k = \frac{1}{4}) = \frac{1}{256}$$

and the probability of

$$(k = \frac{1}{4})(k = \frac{1}{4}) + (k = \frac{1}{4})(k = \frac{1}{4}) = (\frac{1}{256})^2$$

Then,

$$\phi = \left[1 - \frac{2\sqrt{(\frac{1}{4})(\frac{1}{4})(\frac{1}{4})(\frac{1}{4})}}{(\frac{1}{4})(\frac{1}{4}) + (\frac{1}{4})(\frac{1}{4})} \right] 100 = 0$$

$$P(\phi = 0) = (\frac{1}{256})^2 + \sum P(\text{all other paths leading to } \phi = 0)$$

Continuing in this time-honored fashion,* we obtain the distribution of potential percent savings shown in Fig. 7-2. The relative insensitivity of T.C.

* The reader's attention is called to the use of the binomial form on page 118 as a means of obtaining the convolution of this distribution.

to estimation errors is again apparent but it would be even more so if we had not exaggerated the spread of our original k distribution. The purpose in taking an unlikely k distribution was to examine the error response under extremal type conditions. Even so, there is greater than a 65% chance that the potential savings will be less than 10%.

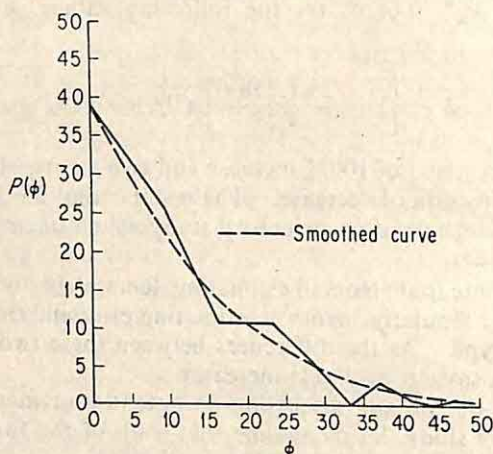


FIG. 7-2 Probability of potential savings percent, ϕ , in a system where four independent k 's are governed by the probability distribution given in the text (----- smoothed curve).

It should be emphasized that the base to which the savings percent, ϕ_k , is applied is not the total average dollar inventory. It is T.C. and the relationship between these two is simply derived. With j items:

$$I = \text{total average dollar inventory} = \sum_j \frac{cx_0}{2} = \sum_j \frac{c}{2} \sqrt{\frac{2zC_r}{cC_c}}$$

$$= \sqrt{\frac{1}{2} \frac{C_r}{C_c}} \sum_j \sqrt{zc}$$

$$\text{T.C.} = \sum_j \sqrt{2zcC_rC_c} = \sqrt{2C_rC_c} \sum_j \sqrt{zc}$$

and
$$\frac{\bar{I}}{\text{T.C.}} = \frac{1}{2C_c} \quad \text{or} \quad 2C_c\bar{I} = \text{T.C.}$$

Therefore, when the carrying cost rate is equal to .5, then the total average dollar inventory equals the total variable cost factor. Carrying cost rates are seldom likely to go above .25. At this latter rate, $\bar{I} = 2\text{T.C.}$ When $C_c = .10$, $\bar{I} = 5\text{T.C.}$ Thus, even a company which has a total average dollar inventory of one million dollars must accept greater than a 65% chance that its savings will be less than $(.10)2C_c(1,000,000)$ and this will be \$50,000 when $C_c = .25$ and \$20,000 when $C_c = .10$. These figures being derived from our extreme error distribution, it is reasonably evident that for average situations management cannot look to ordering quantity improvements for

the sole justification of an intended inventory study. There are other types of savings which we will discuss and when taken together the complete picture—which is required by management—will emerge. However, savings alone do not provide sufficient information to make the kind of rational decisions that we are talking about. Going back to our discussion of a study criterion, we must be in possession of both savings and costs so that they can be compared according to the criterion.

51. VARIANCE OF USAGE AND ITEM GROUPING

Total T.C. is obviously the sum of the individual item T.C.'s. Thus total T.C. = T.C.(Item 1) + T.C.(Item 2) + ... + T.C.(Item N), where the company stocks N items. Part of the function of an inventory study is to

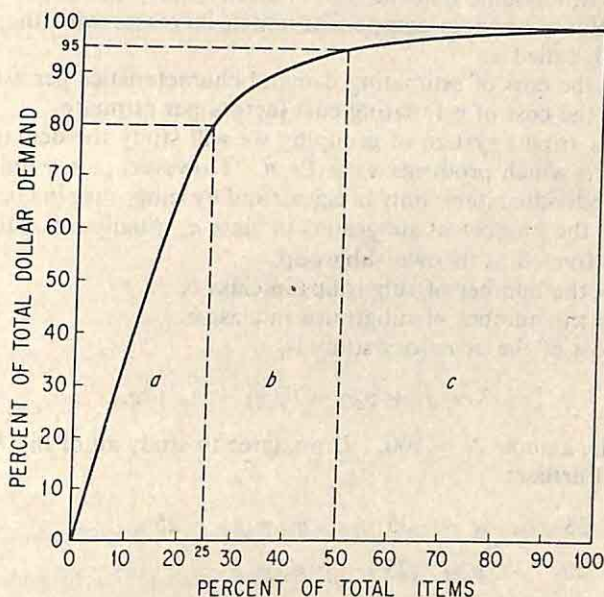


FIG. 7-3 A tri-divisional classification where 25% of the items are responsible for 80% of the total dollar demand.

bring the k factors for each item as close to $k = 1$ as possible. (Here again we assume a $k_1, k_2, \dots, k_5$, and so forth—each k representing the error factor for its appropriate estimate, namely, $C_r, C_c, \dots, C_u$, etc.) A well-known approach to practical inventory solutions is to separate all of the different items which are carried as inventory into two or more classes. The categories are distinguished by levels of dollar usage; i.e., zc . It is generally true that a small number of the total number of items carried contribute overwhelmingly to the total dollar demand of a company. This is graphically portrayed in Fig. 7-3, where the items are arranged in rank order of dollar usage.

The tripart division of this hypothetical curve produces the following results:

25% of the items account for 80% of total dollar demand, these items called *a* have high *zc*. There are n_a such items.

25% of the items account for 15% of total dollar demand, these items called *b* have medium *zc*. There are $n_b - n_a$ such items.

50% of the items account for 5% of total dollar demand, these items called *c* have low *zc*. There are $n_c - n_b$ such items.

In a typical case, the *a* items would be studied individually and treated in great detail; the *b* items would be subclassified into small representative groups; the *c* items would be studied last and would be divided into larger subgroups. It is even possible that the *c* items would not be studied at all.

Now, we will assume that the cost of an inventory study, χ , is some fixed amount *E*, plus a variable component which increases with the number of items studied, called *n*.

Let v_D = the cost of estimating demand characteristics per estimate, and v_c = the cost of estimating cost factors per estimate.

Using our tripart system of grouping we will study the demand for each one of *n* items which produces a cost, $v_D n$. However, cost estimates will be derived for individual items only in class *a* and by subgroups in classes *b* and *c*.

Let g_a = the number of subgroups in class *a*, usually this equals n_a ; i.e., each item is treated as its own subgroup.

g_b = the number of subgroups in class *b*.

g_c = the number of subgroups in class *c*.

Then, the cost of the inventory study is,

$$\chi = E + v_D n + v_c (g_a + g_b + g_c)$$

For example, assume $N = 100$. If we agree to study all of the *N* items then $n = 100$. Further:

$n_a = 25$	$a = .25$	$g_a = n_a = 25$
$n_b - n_a = 25$	$b = .25$	$g_b = 5$
$n_c - n_b = 50$	$c = .50$	$g_c = 5$
$n = 100$	1.00	

$$[g_a = 1, 2, 3, \dots, 25]$$

$$[g_b = (26-30), (31-35), \dots, (46-50)]$$

$$[g_c = (51-60), (61-70), \dots, (91-100)]$$

Letting $E = \$10,000$; $v_D = \$10/\text{item}$; $v_c = \$50/\text{group}$, we obtain,

$$\chi = 10,000 + 10(100) + 50(35) = \$12,750$$

Control over the cost of the inventory study is at least in part vested in the selection and grouping of the items to be studied. We shall, therefore,

investigate the way in which groupings can be chosen as well as the effect of the number of different items carried by the company.

The curve in Fig. 7-3 is of the general type*

$$r_n = \zeta^\beta$$

where $0 \leq \zeta \leq 1$, and $0 \leq \beta \leq 1$.

r = percent of total dollar demand, and $zc(n)$ = dollar demand for the n th item.

$$= \frac{\sum_{n=1}^n zc(n)}{\sum_{n=1}^N zc(n)}$$

(Note: $\sum_{n=1}^N zc(n)$ equals a constant, namely, total dollar usage which we call, T.zc.)

ζ = percent of total items, n/N ,

and β = constant. (When $\beta = 1$ we have a straight line. As β decreases the initial portion of the curve rises more and more steeply.)

From this we obtain:

$$\frac{zc(n)}{T.zc} = r_n - r_{n-1}$$

and

$$zc(n) = \left[\left(\frac{n}{N} \right)^\beta - \left(\frac{n-1}{N} \right)^\beta \right] T.zc$$

$$= \frac{T.zc}{N^\beta} [(n)^\beta - (n-1)^\beta]$$

It will be recalled that T.C. for an item is proportional to the $\sqrt{zc}$. Then the grand total of T.C. will be proportional to $\sum_{n=1}^N \sqrt{zc(n)}$. (We are assuming for the moment that $\sqrt{2C_r C_c}$ is constant for all items. This assumption does not impair our argument and it facilitates understanding.) That is,

$$\text{Total T.C.} = \sqrt{zc(1)}\sqrt{2C_r C_c} + \sqrt{zc(2)}\sqrt{2C_r C_c} + \dots + \sqrt{zc(n)}\sqrt{2C_r C_c}$$

$$= \sqrt{2C_r C_c} \sum_{n=1}^n \sqrt{zc(n)}$$

* Under some circumstances a more precise description of this distribution could be obtained by using the lognormal distribution which is suggested by R. G. Brown, *Statistical Forecasting for Inventory Control* (New York: McGraw-Hill Book Co., Inc., 1959), pp. 197-216.

Expected savings from the inventory study will be proportional to total T.C. and thus proportional to $\sum_{n=1}^{n=n} \sqrt{zc(n)}$. Then,

$$\sqrt{zc(n)} = \sqrt{\frac{T.zc}{N^\beta}} [(n)^\beta - (n-1)^\beta]^{1/2}$$

and
$$\sum_{n=1}^{n=n} \sqrt{zc(n)} = \sqrt{\frac{T.zc}{N^\beta}} \sum_{n=1}^{n=n} [(n)^\beta - (n-1)^\beta]^{1/2}$$

We can now write $E[\psi] = \text{expected savings}$.

$$E[\psi] = \phi(2C_r C_c)^{1/2} (T.zc)^{1/2} \sum_{n=1}^{n=n} [(n)^\beta - (n-1)^\beta]^{1/2} \left(\frac{1}{N}\right)^{\beta/2}$$

where $\phi = \text{expected percent savings}$.

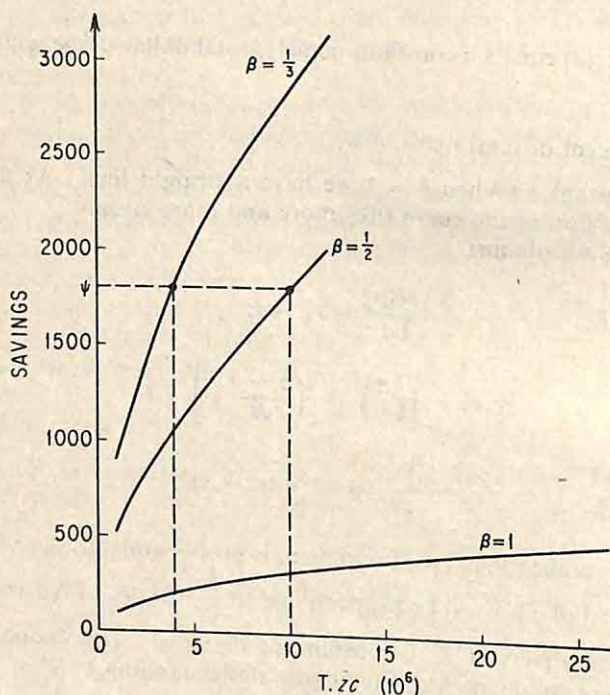


FIG. 7-4 Expected savings as a function of total dollar usage for different levels of β .

$$\left(\begin{array}{l} \text{Savings} = f(T.zc) \\ N = 100 \\ n = 10 \end{array} \right)$$

Examining this last equation we find that savings will increase as $T.zc$ increases. The rate of increase slows down as $T.zc$ gets larger. All of this is what we would expect. And we find in addition that as β decreases (i.e., as the r curve rises more steeply in its initial portion) that the savings level increases. Figure 7-4 illustrates the kind of relationship which holds when

all factors except β and $T.zc$ are held constant. This includes the size of the study which in this case is limited to ten items. Our conclusion is that the more a company's demand is concentrated in a few items the greater the potential savings. Thus, for example, in Fig. 7-4, the company with $\beta = \frac{1}{2}$ must have 2.5 times the $T.zc$ of the company with $\beta = \frac{1}{3}$ to achieve equivalent savings. Consequently, in this regard, the size of the company—as measured by $T.zc$ —is not as influential with respect to savings as is the company's degree of concentration in a few items. It should be remembered that this result does not as yet include the cost of the study. We have been talking

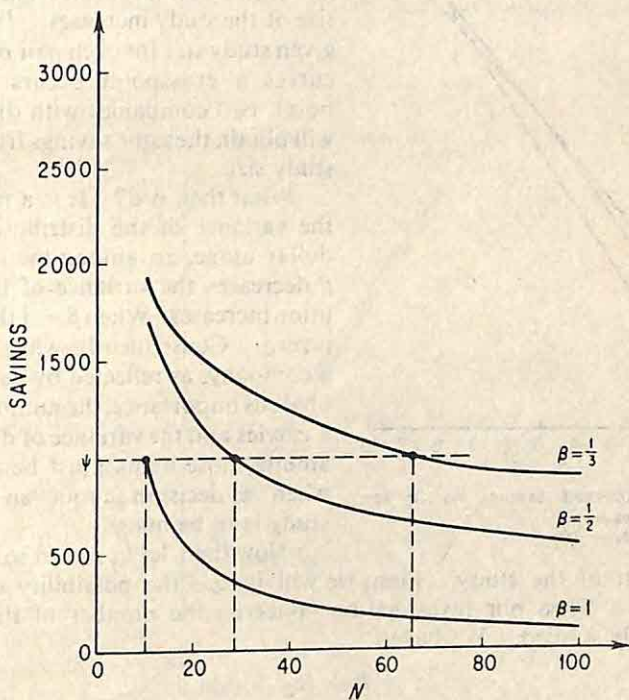


FIG. 7-5 Expected savings as a function of the total number of items carried for different levels of β .

$$\left(\begin{array}{l} \text{Savings} = f(N) \\ T.zc = 10^6 \\ n = 10 \end{array} \right)$$

about gross savings. What we have found is due solely to the dependence of savings on the term, $\sum_{n=1}^{n-n} \sqrt{zc(n)}$.

What about the effect of N on the expected savings? This relationship is graphed in Fig. 7-5 for some representative values. We find that the savings level drops as β increases. The savings level also falls as the company stocks a greater number of items; i.e., with increasing values of N . Again, we point out that the size of the study has been fixed at $n = 10$. We conclude

that the combined effect of β and N make it very difficult indeed to categorize companies as being large and small. To achieve a given level of savings, a company stocking a greater number of items must have a smaller β than a

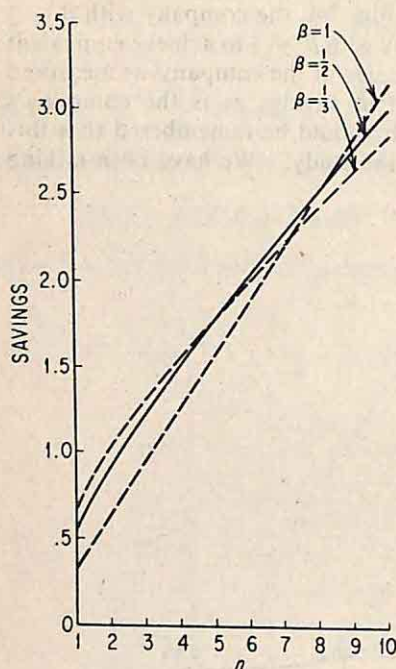


FIG. 7-6 Expected savings vs. n for various values of β . (Fixed $T.zc$, $N = 10$)

company stocking a lesser number of items. Lastly, how does the size of the study affect the expected savings? Figure 7-6 shows how savings grow as the size of the study is expanded. But it also shows that for cases other than $\beta = 1$, the rate of increase slows as the size of the study increases. In fact, at a given study size for each pair of β -valued curves a crosspoint occurs. At this point, two companies with different β 's will obtain the same savings from a given study size.

What then is β ? It is a measure of the variance of the distribution of the dollar usage, zc , among the items. As β decreases the variance of the distribution increases. When $\beta = 1$ the variance is zero. Consequently, while the size of a company, as reflected by its $T.zc$, is of obvious importance, the number of items it carries and the variance of dollar usage among those items must be considered when a decision about an inventory study is to be made.

Now then, let us return to the subject of the cost of the study. First, we will ignore the possibility of creating subgroups. Then our investigation concerns the number of items to be individually studied. We have

$$\chi = E + (v_D + v_c)n$$

$$\text{and } E[\psi] = \phi(2C_r C_c)^{1/2} (T.zc)^{1/2} \sum_{n=1}^{n=N} [(n)^\beta - (n-1)^\beta]^{1/2} \left(\frac{1}{N}\right)^{\beta/2}$$

$$\text{further, } \frac{\text{Required savings}}{\chi} \geq C_c$$

These relations are shown in Fig. 7-7 based on hypothetical data, as follows:

$$E = \$340, v_D = \$50/\text{item}, \text{ and } v_c = \$100/\text{item}. \text{ Therefore,}$$

$$\chi = 340 + 150n$$

$$C_r = \$2/\text{order}, C_c = .10/\text{year}, T.zc = \$10^6/\text{year}, N = 10, \phi = 10\%, \text{ and } \beta = \frac{1}{3}. \text{ Therefore,}$$

$$\text{Expected savings} = E[\psi] = (.10)(.63)(10^3) \sum_{n=1}^n [(n)^{1/3} - (n-1)^{1/3}]^{1/2} (.01)^{1/6}$$

These data have been chosen to permit a simple demonstration with numbers the reader can easily follow. But the principle involved is unchanged when other numbers are used. The crosshatched area in Fig. 7-7 is where:

$$\text{Required savings} \geq xC_c$$

The difference between the expected savings and the base line of required savings is maximum at $n = 5$. But our company, according to its own criterion

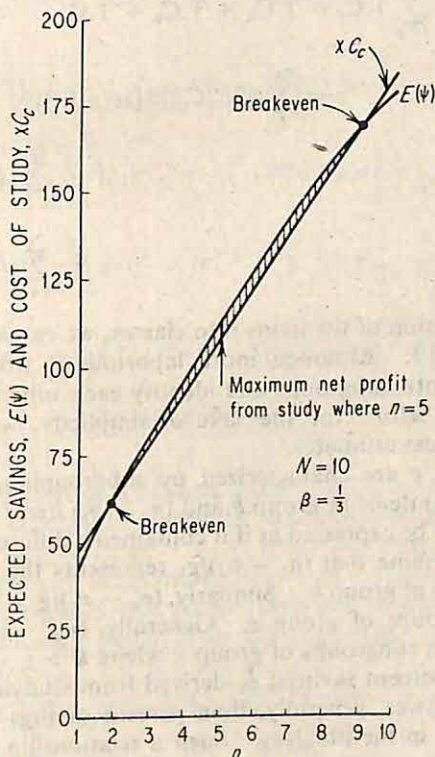


FIG. 7-7 Expected savings, $E(\psi)$ vs. n and cost of study xC_c vs. n . ($N = 10$, $\beta = \frac{1}{3}$)

ion can study any number of items between and including $n = 2$ and $n = 9$. Below $n = 2$ and above $n = 9$, the required savings are greater than the expected savings. But at $n = 5$, the net savings are maximized. Furthermore, it is clear that the optimal point will be unchanged at $n = 5$ no matter what value of C_c is used—even if it represents minimum net loss from an inventory study. On the other hand, the breakeven points will be spaced further apart as C_c is reduced in value and, conversely, closer together as C_c is raised. The maximum value that C_c can be allowed to have for our

example will be when the adjusted cost line is tangent to the β -curve value at $n = 5$. This would be when

$$C_c[340 + 150(5)] = 115.92$$

and

$$C_c \cong .11$$

We will now expand our formulation so that it includes the possibility of using subgroups. The appropriate formulation of the expected savings is quite simply derived. We have three or more classes of items contributing to total T.C. They are counted $1 \dots n_a$, $n_{a+1} \dots n_b$, and $n_{b+1} \dots n_c$.

$$\begin{aligned} \sum_{j=a}^{j=c} T.C_j &= T.C_a + T.C_b + T.C_c \\ &= \sum_1^{n_a} \{(2C_r C_c)_a^{1/2} [zc(n)]^{1/2}\} \\ &\quad + \sum_{n_{a+1}}^{n_b} \{(2C_r C_c)_b^{1/2} [zc(n)]^{1/2}\} \\ &\quad + \sum_{n_{b+1}}^{n_c} \{(2C_r C_c)_c^{1/2} [zc(n)]^{1/2}\} \end{aligned}$$

For any given division of the items into classes, we can associate an average value for $(2C_r C_c)_j^{1/2}$. Although more laborious it would be possible to divide each class into subgroups and identify each subgroup with its appropriate costs. As before, for the sake of simplicity, we will ignore these variations in the cost estimates.

Groups b and c are characterized by subgroupings of different sizes. There are $(n_b - n_a)$ items in group b and $(n_c - n_b)$ items in group c . While each subgroup can be expressed as if it contained a different number of items, it is simpler to assume that $(n_b - n_a)/g_b$ represents the average number of items in subgroups of group b . Similarly, $(n_c - n_b)/g_c$ is the average number of items in subgroups of group c . Generally, $(n_k - n_j)/g_k$ is the average number of items in subgroups of group k where $k > j$.

The expected percent savings, ϕ_k derived from studying subgroups in the k th class will be lower, generally, than percent savings derived from items which are studied in the j th class. Such a relationship might be described by:*

$$\phi_k = f \left[\frac{n_k - n_j}{g_k} \right] = \phi \exp \left[- \frac{1}{w} \left(\frac{n_k - n_j}{g_k} - 1 \right) \right]$$

When $n_k = n_a$, $n_j = 0$ and $g_k = n_a$, subgroups are composed of individual items. We obtain $\phi_a = \phi$ which is the maximum percent savings attainable. Larger subgroups can only obtain fractions of this maximum percent.

This leads us to expressions comparable to those we had before.

* For representational convenience we sometimes denote an exponential relation of the type e^a in the form $\exp [a]$ where e is the natural logarithm base.

$$\chi = E + D[n_a + (n_b - n_a) + (n_c - n_b)] + v_c(g_a + g_b + g_c)$$

and

$$\begin{aligned} E[\psi] = & (T.zc)^{1/2} \left(\frac{1}{N}\right)^{\beta/2} \phi \left\{ (2C_r C_c)_1^{1/2} + (2C_r C_c)_2^{1/2} [(2)^\beta - (1)^\beta]^{1/2} \right. \\ & + \dots + (2C_r C_c)_{n_a}^{1/2} [(n_a)^\beta - (n_a - 1)^\beta]^{1/2} \\ & + (2C_r C_c)_{n_a+1}^{1/2} \sum_{n_{a1}}^{n_{b1}} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_{b1} - n_a - 1) \right] \\ & + (2C_r C_c)_{n_{a1}+1}^{1/2} \sum_{n_{b1}+1}^{n_{b2}} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_{b2} - n_{b1} - 1) \right] \\ & + \dots + (2C_r C_c)_{n_{b2}+1}^{1/2} \sum_{n_{b2}+1}^{n_b} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_b - n_{b2} - 1) \right] \\ & + (2C_r C_c)_{n_b+1}^{1/2} \sum_{n_{b1}+1}^{n_{c1}} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_{c1} - n_b - 1) \right] \\ & + (2C_r C_c)_{n_{c1}+1}^{1/2} \sum_{n_{c1}+1}^{n_{c2}} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_{c2} - n_{c1} - 1) \right] \\ & + \dots + (2C_r C_c)_{n_{c2}+1}^{1/2} \sum_{n_{c2}+1}^{n_c} [(n)^\beta - (n-1)^\beta]^{1/2} \exp \left[-\frac{1}{w} (n_c - n_{c2} - 1) \right] \left. \right\}. \end{aligned}$$

further,
$$\frac{\text{Required savings}}{\chi} \geq C_c$$

We would like to be able to maximize $E[\psi] - \chi C_c$.

Let us begin by taking a "reasonable" division; e.g., for $N = 10$, $n_a = 2$, $n_b = 6$, $n_c = 10$ and $g_a = 2$, $g_b = 2$, $g_c = 1$. We will use the data from the previous example and in addition, let $w = 4$. Further, we will assume, for computational simplicity, that all cost estimates turned out to be the same and that within each class equal subgroup sizes exist. For the cost of the study we can write:

$$\begin{aligned} \chi &= 340 + 50(2 + 4 + 4) + 100(2 + 2 + 1) \\ &= 340 + 500 + 500 = \$1340 \end{aligned}$$

Then

$$\chi C_c = \$1.4$$

Now

$$E[\psi] = (10^6)^{1/2} \left(\frac{1}{10}\right)^{1/6} (10)(2 \cdot 2 \cdot \frac{1}{10}) [1.510 + 1.486e^{-1/4} + 1.160e^{-3/4}] = \$138.$$

and

$$E[\psi] - \chi C_c = +4$$

To find out if we can improve this result we can incrementally change the size of our classes as well as the subgroup sizes. We can also cut the total

size of the study by including less than N items. The table below gives some representative results for different values of the parameters, but we have held the size of the study fixed at 10.

n_a	n_b	n_c	g_a	g_b	g_c	xC_c	$E[\psi]$	$E[\psi] - xC_c$
0	6	10	0	3	1	\$124	\$124	\$0
1	5	10	1	2	1	124	124	0
2	6	10	2	2	1	134	138	+4
4	6	10	4	1	1	144	146	+2
6	10	0	6	2	0	164	168	+4
10	0	0	10	0	0	184	178	-6

On the basis of our limited experimentation it would appear that our original grouping is about as good a result as we can get. It is seen that at least two arrangements of item groupings produce a (+4) net savings and that at least one treatment (all items considered as individuals) yields a (-6) net savings, or loss. Other combinations may produce better or worse results. Furthermore, there is no ready mathematical solution which will yield optimum points. Thus, in the fashion that we have formulated this problem, the cut-and-try method can be used. When all factors including different cost estimates are included, the computational work becomes very formidable. Fast data-processing equipment which would be easy to program would be required if all factors were to be considered. The cost of such computation would be considerable.

Nevertheless, in a sizable inventory study, the strategic grouping of items can contribute significantly to the success of the study and may warrant the kind of analysis which is only feasible with electronic data-processing equipment. Each company has its own particular situation to deal with; i.e., number of items, distribution of dollar usages, total dollar usage, the distribution of ordering and carrying costs among the items, the cost of studying and updating demand, etc. Each company's resolution of the problem would be in both formal and practical terms a unique solution.

52. CONTROLLING THE RESERVE STOCK COSTS OF THE STUDY

We have now seen that in principle—and to some extent, in practice—it is possible to control the cost of an inventory study. It is also possible to estimate study costs and both expected and required savings. But we have limited the scope of our analysis to the area of economic order quantity. As a result, we have understated the size of potential savings which can be derived from an inventory study.

In addition to the losses which result from inefficient ordering policies, we must also consider losses stemming from an incorrect reserve stock level and losses occasioned by forecasting errors. While these losses are related to each other, they are not dependent upon each other. It is completely possible to have perfect forecasts and still be carrying the wrong reserve stock. On the other hand, the determination of reserve stock can exactly reflect the desired

service level while the inability to forecast will create both over- and understocking problems.

We will, once again, attempt to show that the costs of studying these problem areas are at least in part under management's control. And as before, it is reasonable to balance study costs (weighted by the criterion for acceptable gain) against expected savings.

The service level can be determined by judgment. Or, the cost of being out of stock can be estimated and from this figure the appropriate service level is derived. We have already shown in an earlier part of the book* that the service level, α , can be derived from the carrying costs and the understock cost, or conversely, the understock cost can be imputed from the observed or selected service level, and a specified C_c , namely,

$$\alpha = \frac{cC_c}{cC_c + C_u} = \frac{1}{1 + \frac{C_u}{cC_c}}$$

Following our previous pattern, let $k_5 C_u$ be our estimate of the actual understock cost, C_u . Similarly, let $k_6 cC_c$ be our estimate of the actual carrying cost, cC_c . Then our estimated α is α' :

$$\alpha' = \frac{k_6 cC_c}{k_6 cC_c + k_5 C_u} = \frac{1}{1 + \frac{k_5 C_u}{k_6 cC_c}}$$

Let

$$\rho = \frac{C_u}{cC_c} \quad \text{and} \quad \mu = \frac{k_5}{k_6}$$

Whence,

$$\alpha = \frac{1}{1 + \rho} \quad \text{and} \quad \alpha' = \frac{1}{1 + \mu\rho}$$

Therefore,

$$\alpha - \alpha' = \frac{1}{1 + \rho} - \frac{1}{1 + \mu\rho}$$

It will be remembered that:

$$\alpha = \int_{\bar{y}}^{\infty} f(y) dy$$

where $f(y)$ is the distribution of demand during the lead time period and the lower limit y is $K\sigma$ distant from $\bar{y}$. ($\bar{y}$ is the average demand in the lead time period and σ is the standard deviation.)† We observe:

Perfect estimate:

$$(\alpha - \alpha') = 0$$

when

$$\mu = 1$$

* See pages 44-45. (To translate into our present terms: $F(x) = 1 - \alpha$; $K_0 = cC_c$ and $K_u = C_u$.)

† Previously we used s to represent our estimate of the standard deviation. We will assume here that we can obtain the true value, σ although the approach is equally applicable for s .

Overestimate of the protection required:

$$(\alpha - \alpha'_2) > 0$$

when

$$\alpha'_2 = \int_{y+a}^{\infty} f(y) dy \quad \text{and} \quad \mu > 1$$

Underestimate of the protection required:

$$(\alpha - \alpha'_3) < 0$$

when

$$\alpha'_3 = \int_{y-b}^{\infty} f(y) dy \quad \text{and} \quad \mu < 1$$

For almost all shapes of demand distributions which are likely to be encountered (this includes, of course, the exponential, Poisson, and normal distributions) the distribution is monotonically decreasing at the tail. Therefore, for equivalent errors (+ and -) in the service level, from a fixed true α ; i.e., $(\alpha - \alpha'_2) = (\alpha'_3 - \alpha)$, we expect that $b < a$. We can then state that for these equivalent errors (+ and -) in the service level, that:

$$(K_2 - K)\sigma > (K - K_3)\sigma$$

The cost of carrying reserve stock is $(cC_c K_3 \sigma)$ for any estimated α_j . When required protection is overestimated we obtain a potential savings to be made by correcting the estimate of α :

$$\text{Potential savings} = cC_c \sigma (K_2 - K)$$

And when required protection is underestimated by the same amount, we have:

$$\text{Potential savings} = cC_c \sigma (K - K_3)$$

It is apparent that larger savings potential exists for overestimates that can be corrected than for underestimates that can be corrected.*

It is through this relationship that we gain additional control over the cost of the inventory study. Therefore:

(1) Pick items to study with high cost or price, c , and with large variability in their demand. It is no surprise that these two factors emerge from our analysis as being important. But it may be somewhat unexpected that usage, z , has no direct importance here. Of course, the expected usage is reflected in the standard deviation of the exponential distribution and in the variance of the Poisson distribution. The fact remains however, that low demand items with high variance can contribute important savings of the type we are discussing here. Experience with a variety of demand levels indicates that the variability may tend to increase as the demand level decreases. When this is true, the largest potential savings which can be obtained by reducing the reserve stock level are more likely to be concentrated among the low demand items. Further, we may note that high price is also usually associated with reduced demand and that high price is desirable for large potential savings. From this we can deduce that the reserve stock phase of an inventory study concerns at least some of the items which, according to

* This is the reverse of the estimation effect which was observed for ordering costs.

our grouping rules, we would not ordinarily treat in an intensive fashion when investigating optimum order quantities.

Accordingly, every company will have its own particular grouping of items which it expects will contribute the major share of savings. When the order cost grouping does not coincide with the reserve stock grouping an additional study cost must be accepted by the company. In this case, the net savings can be determined for each type of study.

(2) Pick items to study which are more likely to be associated with an overestimate of the protection required. These are the cases where $\mu > 1$. That is, where the error in estimating the stock-out cost is likely to be in the direction of overstating this cost. Or, at the same time we should select items where the error in estimating both price and carrying cost rate tend to underestimate these factors.

At best, these are guide lines. But they formulate conditions with which management can structure both its evaluation study and the full-scale inventory study. They allow a degree of control, based on the most general principles.

53. CONTROLLING THE FORECAST ERROR COSTS OF THE STUDY

The subject of potential savings, to be derived from improved forecasting, requires extensive knowledge of this subject. A basic structure will be developed in Chapter 9. However, for the purpose of rounding out our discussion of controlling study costs we will briefly examine some aspects of the contribution that improved forecasting can make to savings.

Estimates of service level affect the K factor in the expression $(\bar{z}_{L.T.} + K\sigma)$. Here $\bar{z}_{L.T.}$ is the expected demand in the lead time period and $K\sigma$ is the reserve stock level.

Forecasts, on the other hand, affect the value of $\bar{z}_{L.T.}$ and frequently the value of σ . A most obvious case of the latter is when the demand in the lead time period has a Poisson or exponential form. We have already examined the effect of incorrectly estimating demand on the ordering quantity savings. We can assume that the error factor, k_4 , is of the same size when we are estimating demand in the lead time period. Then, we find that, in the Poisson case, the estimated σ , viz. s , is related to the actual σ by the square root of k_4 . Thus:

$$s = \sqrt{k_4} \sqrt{\bar{z}_{L.T.}} = \sigma \sqrt{k_4}$$

For the exponential case we would have,

$$s = k_4 \bar{z}_{L.T.} = \sigma k_4$$

The expected savings in carrying cost to be derived from correcting the demand estimate for a Poisson type distribution when $k_4 > 1$ would be:

$$cC_c K \sigma (\sqrt{k_4} - 1) = cC_c K \sqrt{\bar{z}_{L.T.}} (\sqrt{k_4} - 1)$$

which will always be smaller than the expected savings to be derived from correcting the demand estimate for an exponential type distribution, viz:

$$cC_c K \sigma (k_4 - 1) = cC_c K \bar{z}_{L.T.} (k_4 - 1)$$

Accordingly, forecasting efforts expended on two items which differ only with respect to the shape of their demand distributions in the lead time period will result in basically different savings levels. In this case, the comparison leads us to a preference for items which are exponentially distributed.

Generally speaking, forecasts concern themselves with anticipated demand levels and seldom with the anticipated distribution of those demands in the future. However, when σ is a function of expected demand, consideration should be given to the kinds of distributions which exist and forecasting efforts can be directed to those items which are likely to produce larger savings.

Another factor of importance is the size of the time interval for which the forecast is made. It is clear that as the time period increases, the size of potential forecasting error increases. Therefore, items whose forecasts are prepared infrequently are more likely to produce substantial savings. This fact must then be weighed against the underlying reasons for such infrequent forecasting; e.g., low dollar usage, stable demand pattern, and so forth.

In practice, it is frequently found that forecasting errors account for the greatest per cent of total potential savings. From a representative sample of items a good estimate can be made of the extent to which improved forecasting procedures will pay for the cost of studying a problem. For each item the forecasted demand $F(w_j)$ is compared to the actual demand $D(w_j)$ where w_j is the j th interval of length w .

We let,

$$F(w_j) - D(w_j) = \varepsilon(w_j)$$

and $\varepsilon(w_j)$ is the error term which corresponds to the j th interval of length w . It is simple enough to construct a distribution $f[\varepsilon(w)]$ such that,

$$\int_{-\infty}^{\infty} f[\varepsilon(w)] d\varepsilon(w) = 1$$

and of course, we can obtain the mean value, $\overline{\varepsilon(w)}$ and the variance σ_{ε}^2 .

The fact that the mean value may be zero is of little import. *Forecasting errors do not cancel out.* Figure 7-8 illustrates three different distributions—all with mean zero. These distributions A , B , and C have different amounts of their areas concentrated over different scale points.

When $\varepsilon(w)$ is positive, an excessive amount of stock is carried. When $\varepsilon(w)$ is negative, shortages result. Improvement in forecasts would result in contractions of the distribution on both sides of the zero point. It is better to look at the error distribution in terms of the zero scale point than in terms of the mean value point, since passing through the zero boundary changes the applicable cost. Then,

$$\begin{aligned} \text{Potential savings} = & \xi_1(w, \text{L.T.}) \left\{ cC_c \int_0^{\infty} \varepsilon(w) f[\varepsilon(w)] d\varepsilon(w) \right\} \\ & + \xi_2(w, \text{L.T.}) \left\{ C_u \int_{-\infty}^0 \varepsilon(w) f[\varepsilon(w)] d\varepsilon(w) \right\} \end{aligned}$$

Where, $\xi_1(w, L.T.)$ and $\xi_2(w, L.T.)$ are constants for a given size of w and lead time. These describe the extent to which improved forecasts produce savings for the kind of inventory system which is used, and the extent to which forecasts can be improved by the forecasting method employed. The first term is the potential overstock savings component which occurs when the forecast is greater than the actual demand. The second term is the potential understock component which occurs when the forecast is less than the actual demand. It is quite clear that the greatest potential for savings will be found with high price items having large understock penalties and having as large a σ_e as possible.

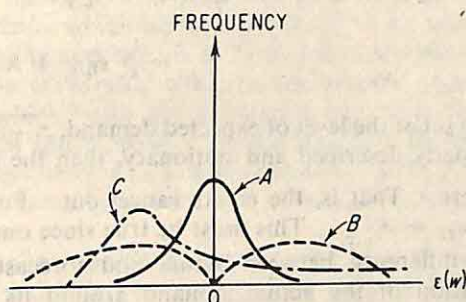


FIG. 7-8 Three different distributions, A, B and C, all with mean forecasting errors equal to zero. $\epsilon(w)$ = size of forecasting error for an interval of length w .

While it is difficult to positively identify in some specific a priori sense those items which have a large error variance, it is quite reasonable to indicate a few of the kinds of considerations which might apply. For example, items which have a distribution of lead times instead of a constant lead time will produce a greater spread in the multiple error distribution. Items which are grouped for forecasting purposes—where the total anticipated demand is then partitioned across the class members (for color, size etc.)—will yield a larger error variance.

A great deal more would be said about forecasting at this point if another even more appropriate niche did not exist for this discussion in Chapter 9. It will be desirable to conclude, however, by highlighting the importance of the error variance. When a forecast of the expected demand in the lead time period is made, it should at least include an estimate of the expected variance of demand in the same period. This is the basis upon which the reorder point is chosen in a reasonable fixed-order quantity system. It is the basis upon which the maximum stock level is set in a reasonable fixed-order cycle system.* Thus, assume that the correct reorder point, P , has been determined by historical data analysis:

$$P = \bar{z}_{L.T.} + K(\sigma_{z_{L.T.}})$$

* The assumption is made that the shape of the distribution is known with some degree of certainty.

where $\overline{z_{L.T.}}$ is the expected demand in the lead time period, and $\sigma_{z_{L.T.}}^2$ is the demand variance in the same period. The reorder point calculation based on the forecast we will call P' .

$$P' = F_{z_{L.T.}} + K(\sigma_{\epsilon_F})$$

$F_{z_{L.T.}}$ is the estimate of demand in the lead time period and $\sigma_{\epsilon_F}^2$ is the variance of the forecast error distribution. Then, summing across a number of forecast periods:

$$\begin{aligned}\sum_i (P' - P) &= \sum_i (F_{z_{L.T.}} - \overline{z_{L.T.}}) + K \sum_i (\sigma_{\epsilon_F} - \sigma_{z_{L.T.}}) \\ &= \sum_i \epsilon_{L.T.} + K \sum_i (\sigma_{\epsilon_F} - \sigma_{z_{L.T.}})\end{aligned}$$

If the forecast is set at the level of expected demand, $\overline{z_{L.T.}}$, and the demand distribution is properly described and stationary, then the first term $\sum_i \epsilon_{L.T.}$ will be equal to zero. That is, the errors cancel out. Further, under the conditions above, $\sigma_{\epsilon_F} = \sigma_{z_{L.T.}}$. This must be true since our forecast is held constant and the difference between actual and forecast values will be precisely the variation of the actual demand around its expected value. Therefore, the sum of the terms $(P' - P)$ over repeated forecasts will equal zero and no additional stock costs will result nor will savings be available. But, a forecast savings could be realized if σ_{ϵ_F} can be made smaller than $\sigma_{z_{L.T.}}$ by the choice of a proper forecasting system which makes use of all possible information.* At the very least, we hope to obtain the zero difference. But, it is perfectly possible to do worse than this. For example, if forecasts have regularly been made by individuals whose interests bias their objectivity or if forecasts have been made by some well-intentioned, but uninformed committee, and so forth, it is entirely possible that a simple statistical analysis of the demand distribution would provide substantial savings. For this reason, it is necessary to update the variance from time to time and the cost of doing this should certainly be included in the inventory study.

54. SELECTING THE SAMPLE FOR THE EVALUATION STUDY

It is apparent that estimates of savings and study costs can be made and that it is important that the right items be included in the study. A company with only a few items can quickly survey the entire scope which the study might take. Companies with such few items are rare birds indeed. Capitalization and sales volume levels are poor guides when it comes to describing the kind of companies which stock few items. Type of industry seems a more appropriate classification. If we consider the grocery, drug store, book store,

* In particular, sequence of demand information where successive values of demand are not independent of each other.

etc., we realize that even very small retail businesses carry an astounding number of items. Small manufacturers producing one simple product should think carefully about their maintenance inventory before they place themselves in the class of organizations with very few items to worry about.

Then, we have ample justification for including so prosaic a subject as the size of the sample to be used in the study. First, how do we draw the sample? At least three different ways suggest themselves. We could take a catalog of all stock numbers and draw a random sample from these. This is an acceptable method as long as we are certain that the sample is in fact a random one. We should get a distribution of cumulative zc which approximates the curve shown in Fig. 7-3. We then analyze this sample for the various characteristics which we have previously been describing.

Another sampling approach is to randomly select days of the year. All items which had a withdrawal ticket on the selected days are listed and the frequency with which items are demanded can then be tabulated. This sampling approach will produce a sample biased in the direction of greater usage. Low usage items will seldom appear. Such a procedure has the advantage of representing activity and transactions instead of the overall inventory situation.

The third sampling approach must be used with caution. Management picks those items which they believe represent the gamut of patterns. Sometimes an even more direct approach is used. Management selects classes of items which conform to their notion of where the important savings can be made. Within each class a sample is chosen—not necessarily by sampling methods—and upon evaluation of a sample an estimated savings is assigned to its class. This last procedure has many pitfalls which are too apparent to dwell upon. However, these methods should not be dismissed out of hand because situations in which they are applicable do exist.

What can we say about the sample size? Let us consider this problem in terms of the a - b - c classes which represent different levels of zc . We draw a small random sample from the catalog. Each of the selected items is then classified according to its annual (or quarterly, etc.) zc , as either a , b , or c . More classes can be used where it seems necessary. As a result of this first small sample we develop probability estimates p_a , p_b , and p_c —that is, the probability of drawing an a , b , or c item.

If this first small sample is sufficient, we must have confidence that the estimates p_a , p_b , and p_c each fall within some limited range, such as $p_a \pm \Delta$. The range $\pm \Delta$ is the required precision interval where Δ is expressed as a decimal fraction. We will call this interval the level of accuracy required. Now, we must also choose a level of confidence such that this is the confidence we require that the actual values of p_a , p_b , and p_c lie within the specified range of accuracy.

Using the parameters of the binomial distribution we can set down the following relationship:

$$n_j = \left(\frac{K}{\Delta} \right)^2 p_j (1 - p_j)$$

where n_j = the indicated sample size for the j th group,

K = the number of standard deviations which are needed to produce a given confidence level,

Δ = the decimal fraction indicating the specified range of accuracy,

p_j = the observed frequency of the j th group expressed as a decimal fraction of the total number of items observed.

To illustrate the use of this simple formula we will require an accuracy of $\pm 3\%$, with a confidence level of 95.4%. Assume that our first selection of 50 items resulted in $p_a = .60$, $p_b = .30$, and $p_c = .10$. We associate $K = 2$ with a 95.4% confidence level. So,

$$n_a = (2/.03)^2(.60)(.40) = 1067$$

$$n_b = (2/.03)^2(.30)(.70) = 933$$

$$n_c = (2/.03)^2(.10)(.90) = 400$$

It is obvious that at this stage, a much larger sample is indicated where the largest requirements, namely, $n_a = 1067$ dominates the other classes. However, it is hardly sensible to select another 1017 items before checking again to learn what modifications in sample size are indicated.

We will suppose that when the sample consists of 200 items there are then 150 items in class a , 30 items in b , and 20 items in c . Once again we check to determine the indicated sample size. This time for $p_a = .75 \pm .03$, $p_b = .15 \pm .03$, and $p_c = .10 \pm .03$. We find $n_a = 833$, $n_b = 566$, and $n_c = 400$. Continuing in this way, the actual sample taken eventually coincides with or exceeds the largest requirements derived from the formula.

We may note in passing that the required sample size will be largest for specified levels of accuracy and confidence when $p_j \approx .50$. Therefore, it is advisable to name the classes a , b , and c in a way which will avoid giving anyone of them half or nearly half of the total item representation. This philosophy is in keeping with the kind of curve represented by Fig. 7-3.

It is also desirable to mention that the method demonstrated here has usefulness as well in determining sample sizes during the actual inventory study. And it can be employed to determine the required size of a Monte Carlo or simulation run.

A more stringent specification of accuracy is obtained when the range is stated as being some fixed percent of the "true" value. Then we have:

$$p_j(1 + \Delta) = p_j \pm K \sqrt{\frac{p_j(1 - p_j)}{n_j}}$$

which gives us:

$$n_j = \left(\frac{K}{\Delta}\right)^2 \left(\frac{1 - p_j}{p_j}\right)$$

Using the same specifications as were previously set, we obtain these results:

$$n_a = 2964, \quad n_b = 10,367, \quad n_c = 40,000$$

Now the c group dominates the sample size. The rank order of the group sample sizes is reversed. The result is entirely in keeping with the newly specified ranges, namely:

$$p_a = .60 \pm .018, \quad p_b = .30 \pm .009, \quad p_c = .10 \pm .003$$

With this type of specification of accuracy, the required sample size grows smaller as the size of the group grows larger. It is not reasonable to permit the extremely narrow range of the small p groups to dominate the sample size. Therefore, we should either specify larger ranges for low p values or else use the first formulation which was given above.

When the total number of items, N , carried by the company is small, neither of the relationships above will properly serve the intended function. This is due to the fact that we are sampling the total stock of items without replacement, a condition which falls outside of the binomial requirements. Usually, when N is small, there would not be a sampling problem but for those cases where there is, we give the following equation:

$$p_j + \Delta =$$

$$\frac{\left\{ p_j n_j + \frac{K^2}{2} \left(\frac{N - n_j}{N - 1} \right) \pm \sqrt{\left[p_j n_j + \frac{K^2}{2} \left(\frac{N - n_j}{N - 1} \right) \right]^2 - \left[n_j + K^2 \left(\frac{N - n_j}{N - 1} \right) n_j p_j^2 \right]} \right\}}{\left[n_j + K^2 \left(\frac{N - n_j}{N - 1} \right) \right]}$$

This formula parallels our first relationship for sample size, since the precision interval is stated as a fixed range and not as a percentage of the "true" value. When N is large, relative to n_j :

$$\left(\frac{N - n_j}{N - 1} \right) \rightarrow 1$$

and so we can examine without difficulty, the effect of decreasing values of N on the confidence level of a specified precision interval and sample size. We let

$$K = c \left(\frac{N - 1}{N - n_j} \right)^{1/2}$$

where c = constant.

When N is large relative to n_j , $K = c$. However, as N gets smaller with respect to the sample size n_j , K grows rapidly in size. Since K is the effective multiplier of the standard deviation, its size determines the confidence level. Thus, when we specified a 95.4% confidence level we observed that $K = 2$. From our analysis, it is clear that the confidence level increases as N decreases -holding the sample size and the precision interval constant.

Figure 7-9 shows how the confidence level decreases with increasing N for our three sample sizes, n_a , n_b , and n_c when the precision interval is fixed. We observe that in general if $N/n_j \geq 10$, a specified confidence level will not be too sorely abused. Below this limit, we will be oversampling, and to avoid this situation we can use the more complicated formulation with

fixed K , Δ , N , and our observed p_j . We then solve iteratively for a new n_j .

A word should be said about the classification scheme for groups a , b , c , ..., etc. It is *not always desirable* to group items by zc level. Sometimes we might want to classify by variance of the demand, by forecasting character (such as seasonal and nonseasonal items), by lead time character, by differing service level and so forth. All of these classifications are perfectly legitimate. The groupings should be designed so that the essential cost and potential

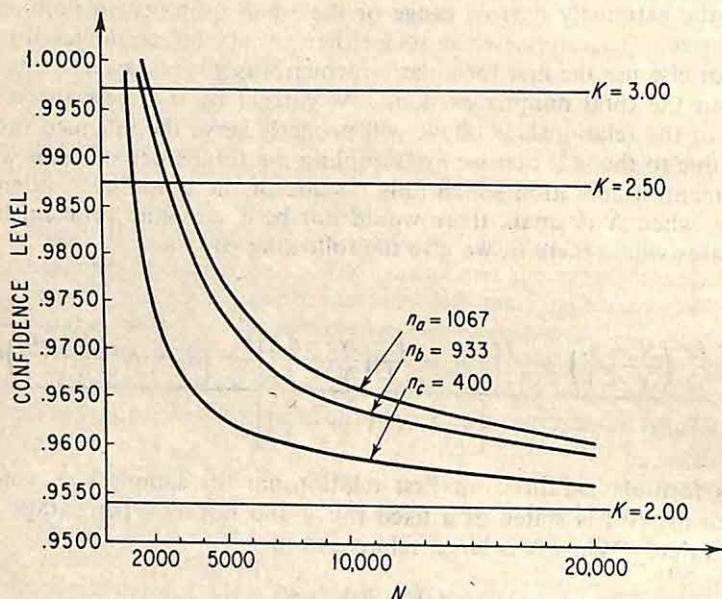


FIG. 7-9 The confidence level increases as the number of items carried by the company, N , decreases. The precision interval is fixed at $\pm .03$.

saving factors are laid bare for analysis. Combinations of factors at different levels can be used. Such groups should be designed to be mutually exclusive of each other and taken as a whole they should be all inclusive of the total number of items carried.

Another aspect of the sampling problem which strongly influences the cost of the study is the amount of past history which should be studied for each item. In general, the amount of data which is to be used in the evaluation study should be sufficient to reveal seasonal variations and trends—normally a two-year period will do. Further, it should isolate and separate portions of the data which may be lumped under a single stock number. For example, the item might have been redesigned or received special promotion at some time during the interval which the data covers.

Our discussion of sample size reveals, once again, that in an area where intuition and judgement are so frequently applied, it is totally reasonable to formalize the problem so that intelligent decisions can be made which can at least guide management's intuition. Even the intuitive evaluation of an

inventory system can proceed on a more rational basis if the executive does not rely on a visit to the stock room and a careful flipping through the stock cards. If the right (representative) set of stock cards are presented to the executive, he can at least spare himself the effort required to generalize from some poorly drawn sample.

55. EXHAUSTION OF STUDY FUNDS

We have reached a point where we can categorize the intuitive approach as a form of gambling. There are so many factors to consider that an evaluation study, if pursued to its logical end, represents a large-scale study in itself. But if no formal evaluation is used, then the decision to undertake a study is based more on the fact that management is willing to risk a sum of money with the hope that an adequate return will be forthcoming. The picture is not quite as black as we appear to be painting it, because we assume that management prefers to bet on situations where the odds are in its favor. Let us state this in a somewhat different way. The intuitions and judgement of management used in place of analytical evaluation is presumed to produce a range of possible returns for a given level of investment in a study. All of those returns are considered possible. By intuition, a given probability is associated with each possible return. The description we have just given fits closely a number of situations in which a gambler might find himself thinking in much the same way. The connotation of gambling is one which an executive might prefer to avoid. This is somewhat misleading since the difference between risk management and gambling is primarily semantic. In many respects the gambler's performance is the analog of a series of executive decisions which can lead ultimately to executive or corporate ruin.*

Interesting results can be derived from this analogy. For one thing, we find that companies with low capitalization must assume disproportionately large risks of exhausting their supply of research and study funds—as compared to companies with higher capitalization. To demonstrate this point and develop several others, we will make use of the classical problem of the gambler's ruin.†

For our purposes we will interpret ruin as the total exhaustion of a company's capital which can be employed in product, process, and operations research. We are including inventory studies in the last of these three categories. The analysis can begin by providing some definitions of symbols which will be used thereafter. Let,

κ = available capital as defined above,

χ = cost of the inventory study,

$\zeta = \kappa/\chi$ = the company's available capital as a multiple of study cost.

* D. W. Miller and M. K. Starr, *Executive Decisions and Operations Research* (Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1960), pp. 366–375.

† See J. V. Uspensky, *Introduction to Mathematical Probability* (New York: McGraw-Hill Book Co., Inc., 1937), Chap. 8. Also see William Feller, *An Introduction to Probability Theory and its Application* (New York: John Wiley & Sons, 1950), pp. 297–303.

We know that, $\zeta(\text{large company}) > \zeta(\text{small company})$, since the fixed costs of an inventory study form a substantial portion of total study costs.

On its available capital, κ , the company expects to get a return of κC_c , where C_c , as before, is the carrying rate criterion. Thus, at the end of a year the company's objective is to have achieved the quantity a ,

$$a = \frac{\kappa + \kappa C_c}{\chi} = \zeta(1 + C_c)$$

where a is expressed in multiples of χ . If a is achieved then the company has succeeded in preserving its capital and adding to it an amount which is in keeping with its own criterion.

Lastly, we must explore the probability distribution of savings which might be obtained as a result of our inventory study. We choose, m , as the largest feasible savings. We let m be expressed as a multiple of χ . Therefore, when $m = 1$, our savings exactly equal our cost and we have broken even. Obviously, no executive is going to underwrite an inventory study where the largest likely return is m equal to one. We will suppose that the probability distribution of savings is exponentially distributed with a mean of $(1 + C_c)$. Because of its exponential form, the standard deviation is equal to the mean. We will advance m to a position 3σ from the mean in a positive direction. Thus,

$$m = \text{Mean} + 3\sigma = (1 + C_c) + 3(1 + C_c) = 4(1 + C_c)$$

The lower limit of our distribution has been chosen equal to zero. This is equivalent to losing χ dollars since the entire cost of the study has been of no avail; i.e., no savings benefit, whatever. We would not ordinarily expect to encounter anything worse than this, although it is possible that a study might create such confusion and lowering of morale that a loss greater than χ would result. Such possibilities can be taken into account by modifying the ruin formulation so that it correctly represents whatever conditions prevail. For our assumptions we can write two equations* as follows, with roots 1 and Δ :

$$u_\zeta = \frac{\Delta^a - \Delta^\zeta}{\Delta^a - \Delta}$$

$$u_\zeta = \frac{\Delta^{a+m-1} - \Delta^\zeta}{\Delta^{a+m-1} - 1}$$

where u_ζ = the probability of exhausting the available capital.

Substituting the expressions we have derived for a , and m , we obtain:

$$u_\zeta = \frac{\Delta^{\zeta(1+C_c)} - \Delta^\zeta}{\Delta^{\zeta(1+C_c)} - \Delta}$$

$$u_\zeta = \frac{\Delta^{\zeta(1+C_c)+3+4C_c} - \Delta^\zeta}{\Delta^{\zeta(1+C_c)+3+4C_c} - 1}$$

* William Feller, work previously cited.

The solution of these equations requires that we find the second root, $\Delta \neq 1$. We can do this by balancing the right- and left-hand sides of the equation:

$$(1 - \Delta) = \Delta^{\zeta C_c} (1 + \Delta^{\zeta+3+4C_c} - \Delta^{\zeta} - \Delta^{4+4C_c})$$

Substituting back into either equation yields u_{ζ} .

Figures 7-10 and 7-11 illustrate the way in which the probability of exhaustion of capital changes with ζ and with C_c .

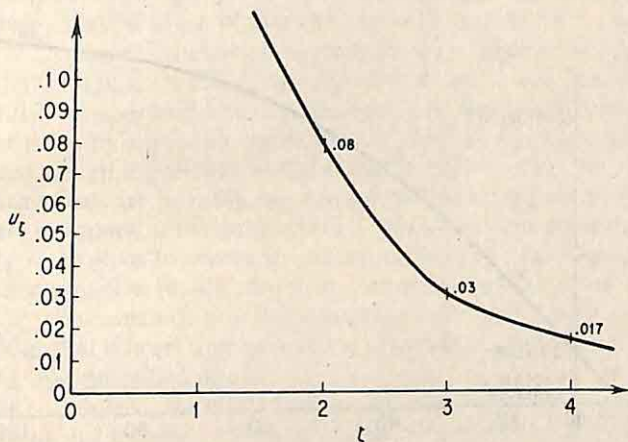


FIG. 7-10 Probability of ruin as a function of the company's capital-cost multiple, ($C_c = .10$).

(1) As the ratio of available capital to study cost increases, the probability of exhausting the available capital decreases. That is, as ζ increases, u_{ζ} decreases. The result, in itself, merely serves to confirm what we expected. However, the decrease is very fast initially and slows up as ζ becomes larger. This means that companies with small research and study budgets run large risks of exhausting their funds. At the same time, reductions in ruin probabilities become increasingly difficult to obtain as the capital fund increases. A completely literal interpretation of the numbers shown in the figures is unreasonable. However, they suffice to give us an approximate idea of the risks involved. We may note from Fig. 7-10 that if $\zeta = 1$; i.e., all available capital is allocated in a single study—the probability of ruin goes up to about 46%. This still assures an edge for success but not a very good one. When only half of the available funds are allocated to a single study, the probability of ruin drops to the neighborhood of 8%. Thus by doubling the capital to cost ratio, we divide the ruin probability by more than five. This clearly indicates the advantage which companies with large capital resources have. It also provides a very crude rule of thumb with respect to which companies should consider intensive study of their inventory system. Namely companies whose capital is less than two and a half, $u_{\zeta} \approx .05$, or three, $u_{\zeta} \approx .03$,

times their estimated study costs are exposing themselves to undue risk. Of course, if the distribution of probable savings is more favourable, the cut-off level will be raised.

(2) The opposite effect can be observed when the criterion for undertaking a study, C_c , is raised. This is apparent from Fig. 7-11. It will be remembered that the distribution of potential savings was stated in terms of C_c . Thus, studies with a higher C_c were expected to have a better promise of return.

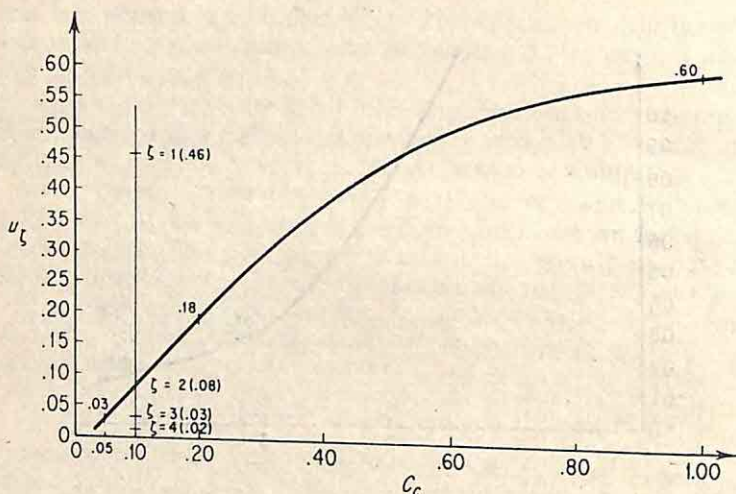


FIG. 7-11 Probability of ruin as a function of the carrying rate, ($z = 2$).

In spite of this fact, a larger carrying rate increases the probability of ruin. Why is this so? It is because the company cannot just skim the cream. They must study the inventory system in greater depth in order to obtain the return which their more demanding criterion requires. This implies that the study will take longer and will involve greater risk of frittering away the initial savings which are obtained in pursuit of additional savings from what is essentially a poorer grade of ore. But two things can be noticed about the $u_z \sim C_c$ effect. First, the negative response of u_z to increasing C_c diminishes as C_c approaches one. Within a reasonable range of values for C_c , say .05 to .30, the relationship shown is almost linear, and u_z increases at about the same rate as does C_c . This observation underscores the importance of the company's criterion in determining the long-range success or failure of the company's research and study program. Second, for the value $C_c = .10$, the ruin probability is indicated for different values of z (Fig. 7-11). Naturally, the effect of changing the size of z on u_z is the same as that observed in Fig. 7-10. But here we have the opportunity of comparing the z effect with the C_c effect. We observe that decreasing z from two to one results in a ruin probability equivalent to increasing C_c from .10 to about .52. At least in this sensitive region of low z , the z effect tends to dominate the situation.

56. CONCLUSION

At the beginning of this chapter, we asked why management is prone to ignore a formal analysis of certain systemic costs which affect many of their decisions. Although, up to this point, we have not discussed all systemic costs which fit this category, nevertheless, the answer to the question is sufficiently apparent in terms of the preceding material.

Potential savings are the result of errors in estimation, inefficient decision procedures, and processing routines. The present chapter has dealt primarily with the first point—although this is inextricably bound up with the second point. That is, the behavior of a company with respect to its inventory—whether it is “rational” behavior or not—implies specific estimates in a rational system. Thus, in effect, we can transform a conglomeration of decisions which are sporadic, changeable, and even whimsical into a system of consistent rules by imputing the values of the parameters which would have produced something similar to the actual behavior. To the extent that such transformations are possible, we can derive the savings to be gained by more accurate estimates of the parameters. We know that potential savings are relatively insensitive to errors in the estimation of ordering quantities. They are more sensitive to the errors in estimating service level and forecasting errors. We conclude that it is possible to evaluate potential savings—but logic tells us that it is not always feasible to conduct a full-scale evaluation study such as that described above.

Then, what is needed is the kind of rules of thumb which we have developed and discussed in this chapter. Such guidelines are based on the number of items carried, the variance of dollar usage among the items, the variance of usage over time for each item, and so forth. When this information is assembled on the basis of a properly designed sample, the need for a full-scale evaluation study is eliminated.

Furthermore, a great deal of control exists over the study costs and the probability of ruin. A company with fixed available resources for research and analysis can equate their study criterion with the probability of ruin and thereby determine at the outset whether or not to undertake a study. They can reduce the ruin probability by lowering the criterion value. But why should they do this if the risk of investing the capital outside or using it for other research projects is lower while the return on investment is equivalent or better?

Another control exists in fixing the size of the study. Then, for a fixed level of χ , it is possible to decide whether to study ordering quantity, service level, forecasts, or the data-processing system. Perhaps a combination of these is in order. Control exists over cost and savings within each area by means of grouping and subgrouping and by means of proper item selection.

So, we conclude that the answer to our original question is that management has good reason for tending to ignore complex systemic cost estimates only so long as the area is underdeveloped. Few companies, if any, would see the advantage of spending their study funds developing generalized analytic

frameworks such as we have been doing here. However, since the area of systemic costs has been receiving increasing attention, it is likely that management will utilize these decision-making aids with increasing frequency. Ultimately, the decision to undertake an inventory study will be weighed against the advantages to be derived from a new machine, research on a new product, or an extra dividend for stockholders. And in addition, more efficient studies will result.

Within the limitations of a totally general discussion, we have shown that a great deal of control exists in the hands of the decision-maker with respect to the net savings which can be derived from an inventory study. In spite of this, management is still in a predicament. First of all, the area of control is incredibly obscured by a multiplicity of parameters and variables. Larger companies have access to the kind of personnel that can cope with such systems. But the smaller companies, which, as we have shown, must be more precise in reaching their decisions, are usually unable to draw upon the company staff for the purposes of a formal evaluation study. The expense of external assistance for evaluation purposes may not be easily justified. Even if the executives of the company can undertake a reasonable evaluation and produce a meaningful estimate of optimal net savings—this procedure entails its own cost which the smaller company may not be in a position to underwrite. So the odds are that the smaller companies will continue to use common sense, comparable industry experiences, and knowledge of how things seem to be going. This may be unfortunate but there does not appear to be an easy way out of the dilemma.

Even with large companies the question must be answered, how much pre-study or evaluation should we use? The cost of defining and obtaining an adequate sample is a pre-study cost. The cost of determining an optimum cost for an evaluation study is a pre-pre-study cost. And so we go from pre-study to pre-pre-study and the chain stretches backwards to oblivion. Decision chains of deciding to decide to decide, etc., are not unfamiliar to management. That is why there comes a point—different for each case—when management is ready to gamble on its intuition. The point should be fixed by the fact that the expected loss from bad judgement is small and easily tolerated by the company. The ruin analysis offers at least a guide for determining when such a gamble is justified. And it should be remembered that the decision not to study a system—if incorrect—can be as bad, if not worse—in terms of opportunity costs than the incorrect decision to study a system.

The importance of a reasonable evaluation procedure, whatever form it takes, cannot be overstated. Companies with sound, although nonformal, inventory procedures may derive only marginal benefits from a full-blown inventory study. It may be that only the forecasting area should be studied, or it may be that only *a* class items should be studied. Abstracting from all that we have previously discussed: the larger the company's sales volume, the more diversified the line that is carried, the greater the variety of maintenance stock, the more complex the production line, the greater the span of executive control, the more likely that an inventory study is warranted.

PROBLEMS

1. Three executives at a meeting discuss their individual criteria for accepting a contemplated inventory study. The first requires that $\psi/(x_1 + x_2) \geq .20$. The second indicates his preference for $\psi p_5/(x_1 + p_5 x_2) \geq 1.00$. The third executive refuses to accept either basis and offers his choice, $\mu\psi/(x_1 + \mu x_2) \geq 2.00$. (The symbols are the same as those in the text, and management uses the value $C_c = .06$.) If $x_1 = \$50,000$; $x_2 = \$30,000$ per year; and the expected savings, $\psi = \$40,000$ per year, whose criterion would accept the study? Comment on the results.

2. How do economies of scale affect the likelihood of small and large companies undertaking inventory studies? How does this affect companies with subsidiaries?

3. If estimation errors of +20%, -30%, and +50% are made in the forecast of demand, ordering cost, and carrying cost, respectively, what distortion could be introduced in the estimate of average price in order to balance the errors with respect to the total variable costs? Is the result intuitively obvious?

4. Our company wishes to investigate how many items should be studied before embarking on an inventory program. Each item is to be treated individually. The relevant data are as follows:

N = total number of items stocked = 625 items,

E = fixed study costs = \$2000,

v_D = cost of estimating demand characteristics per estimate = \$60 per item,

v_c = cost of estimating cost factors per item = \$100 per item,

C_r = cost per order = \$20.00 per order,

C_c = carrying cost rate per year = .10 per year,

$T.z.c$ = total yearly dollar demand = \$4,000,000 per year,

ϕ = expected percent savings = .10,

β = measure of dollar-demand variance among items = 1/2.

(a) What cost equation describes the study?

(b) Write the equation for expected dollar savings per year.

(c) Using the criterion:

$$\frac{\text{Required savings}}{\text{Study cost}} \geq C_c$$

find the acceptable minimum number of items to be studied.

5. Compare the reserve stock savings for two items, both normally distributed with mean of 40 and standard deviation of 5. The protection level for item A has been overestimated by the same amount that the protection level for item B has been underestimated, namely, .025. Other data are as follows:

	PRICE	C_c	α'
A	\$8	.10	.050
B	\$4	.10	.050

Comment on the result.

6. An evaluation study is being undertaken prior to embarking on a full-scale inventory study. An *a-b-c* division of items is used and the first sample of items results in the following observed frequencies.

$$a = 40 \text{ items}, \quad b = 20 \text{ items}, \quad c = 40 \text{ items}$$

We require an accuracy of $\pm .05$ with a confidence level of 95.4%.

(a) How large a sample is indicated at this point? Comment on the advisability of using this particular definition of the groups.

(b) If the company stocks 10,000 different items, will any change in the planned sample size be called for?

7. Compare the ruin probability of two companies equivalent in every way except that one company insists on the criterion for undertaking an inventory study that $C_c = .30$, while the other sets $C_c = .05$ as their basis. Comment.

the inventory study: design of decision procedures

As we have previously pointed out, the cost of the evaluation study is not recoverable if the result of this study is the decision to go no further. Our statement must, however, be tempered by the fact that certain positive benefits may accrue during the course of the evaluation. But our primary interest is directed toward the major savings which can only be realized as a result of the actual inventory study. Consequently, we begin this chapter with the assumption that a positive result has been obtained from the evaluation study. Then, a logical primary procedure is to employ systems analysis which is a powerful, analytical means of organizing the problem area. Systems analysis is by no means restricted to inventory studies. It is generally applicable when an intensive effort is undertaken to improve any complex system.

57. SYSTEMS ANALYSIS

Our method provides a means for investigating whole systems and integrated procedures. It is fair to say that the rapid growth in the use of this approach for inventory studies can be attributed to the advantages of designing inventory procedures which could be handled by a computer. Various discussions of the application of systems analysis to inventory problems have appeared in recent literature but a good part of these references appear in articles primarily concerned with computers and electronic data-processing, or EDP. The importance of systems theory to inventory analysis should not be predicated on the use of a computer. The benefits which can be derived even when electronic data-processing is not involved are substantial.

As we proceed, we are going to concern ourselves with several different levels of systems analysis beginning at first with overall system views and then moving into finer detail until we are finally dealing with the ultimate detail required by the computer.

Systems analysis has been known and used in other fields for a long time. Engineers have subscribed to its thinking and have applied its methods in many engineering areas. Only of recent vintage however is the broad application of systems theory to business problems of many kinds including inventory.

There is nothing startling, revolutionary, or mysterious about systems analysis. The approach is best epitomized by the word, "systematic." Art and science have many affinities. The paths appear to diverge, however, when the question of systematic elucidation of detail arises. The art of executive decision-making is similarly in conflict with the scientific management approach. While the conflict is mostly illusory, it nevertheless is responsible for a reluctance on the part of many people—executives are not excepted—to be entirely systematic when such behavior is indicated. Whatever the underlying reasons may be, the fact is that many people tend to avoid the burden of tracing through seemingly endless labyrinths of detail. This fact has undoubtedly been a major deterrent to the effective use of system analysis in business, institutional, and government operations.

We can all recognize that Herbert Simon's principle of bounded rationality* is a reasonable description of man's approach to complex problems. Namely, man's rationality operates within the framework of a simplified model of a real situation. Man's brain is not capable of assimilating and considering all of the factors which might be relevant to a given problem area. The problem-solver or decision-maker cannot conceive the totality of a complex system and he cannot employ all cogent factors in his cerebral payoff matrix. So he discards most factors. He reaches his conclusions on the basis of a few factors—somehow carefully chosen because they seemed most important.

There are problems, however, which when tackled in this manner produce decidedly poor results. Usually these are problems in which the entire system must be viewed as a whole in order to design the components. For example, weapon systems, guided missiles, and automatic factory operations are just too complex to be satisfied by bounded rational solutions. There may be thousands of components each capable of destroying the overall effectiveness of the system. A natural question is, why don't we tackle each component as a subsystem? Such subsystems can be reduced to the point where the bounded capacity of the decision-maker can effectively cope with the limited degree of complexity of the subsystem. And the answer is that for some systems we can do just that. But for most of the systems which interest the executive we cannot. Inventory systems are of the latter type. Why is this so? There are a number of reasons and we will briefly explain several of them.

* Herbert A. Simon, *Models of Man, Social and Rational* (New York: John Wiley & Sons, Inc., 1957), pp. 196-206.

58. THE MANY ITEM EFFECT

One of the major reasons is that most companies stock more than one item. We have previously mentioned the effects of grouping items with respect to ordering costs, delivery rates, and so forth. It is now desirable to expand upon these notions. As we shall show, many kinds of dependencies may exist among items—a fact which can be ignored only at the decision-maker's peril. The *many item effect* is felt in a variety of ways. The study cost will increase incrementally in a monotonic fashion as more items are included in the study. Sharp increases in cost and other discontinuities will be encountered when a new model type is required. This includes the cases in which different forecasting routines and different data-processing programs are employed. Our discussion in the previous chapter with respect to grouping principles emphasizes the fact that dependencies link the characteristics of the items with the expected gain from the study. This inhibits treating each item as a separate component.

Another consideration of significance in a realistic situation is that the company's capital is not unlimited. The fact that many items are carried influences our decisions with respect to how we should employ the capital which is available for inventory investment. When a capital budget exists, we must choose between items in a very practical way.* Consequently, it is impossible to consider the items as individual components, each of which can be separately optimized. Let us illustrate this phenomenon by demonstrating our point with two items. The same conclusion will be obtained if many items are approached in the same way. The method we are using is another form of the dynamic programming† technique which we have previously used in this book.

We have two items P_1 and P_2 with the following parameters:

	c	C_c	C_r	z	x_0	$cx_0/2$
P_1	\$2	.10	\$4	10	20	\$20
P_2	5	.10	4	16	16	40
						\$60

$cx_0/2$ = optimal average dollar inventory. If we can afford a total average dollar inventory of \$60, all well and good. We simply follow the indicated optimal policy of ordering P_1 and P_2 in quantities of 20 and 16 units respectively. But let's say we have only \$40. What should we do? Applying our systems approach we recognize that the elements (items) can no longer be considered to be independent. The matrix in Fig. 8-1 shows

* The reader will remember our development of the unknown cost system (pages 93-104). The dynamic program which we are about to use will provide the same kind of result as the unknown cost system. However, the dynamic programming method provides a systems generalization which is very flexible and applicable to a great variety of situations not inherent in the unknown cost system.

† Richard E. Bellman, *Dynamic Programming* (Princeton, N.J.: Princeton University Press, 1957).

precisely what happens for all possible combinations of total average dollar inventory investment for these two items in the range \$10 to \$90. The entries in the matrix represent total variable costs, i.e., $\sum_j T.C_j$. The intersection of the row and column marked by asterisks is the optimal or lowest value of $(T.C_1 + T.C_2)$. No matter how we expand the matrix, it will always be the lowest value. Note that this lowest value falls on the diagonal marked \$60. It is the ordering policy which was derived above by obtaining the optimum values individually for each of the items considered. Specifically, for P_1 the average dollar inventory is indicated to be \$20, and for P_2 it is \$40.

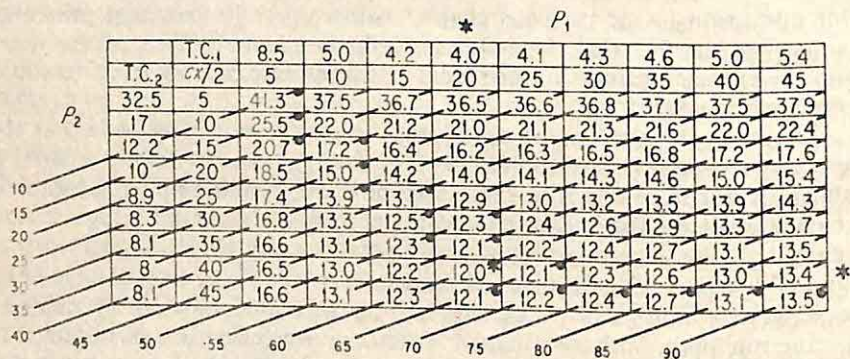


FIG. 8-1 Matrix of total variable costs for two items, P_1 and P_2 , where the total average dollar inventory can take on values in the range \$10-\$90.

Each diagonal of the matrix includes all combinations of the average dollar inventory investments that can be made in P_1 and P_2 and which sum to a fixed amount. Consider the line marked \$40. That is how much money we said we have available for our average dollar inventory investment. The \$40 total results from combinations of \$35 invested in P_2 and \$5 invested in P_1 , \$30 invested in P_2 and \$10 invested in P_1 , \$25 invested in P_2 and \$15 invested in P_1 , and so forth. Specification for each item of the average inventory investment, in addition to the parameters given above, is sufficient to determine $T.C_1$, $T.C_2$ and $\sum_j T.C_j$. Thus,

$$T.C. = \frac{cx C_c}{2} + \frac{z}{x} C_r$$

An example of a calculation is shown below.

	$cx/2$	cx	x	$cx C_c/2$	$z C_r/x$	T.C.
P_1	30	60	30	3.0	1.3	4.3
P_2	10	20	4	1.0	16.0	17.0
						$\sum_j T.C_j = 21.3$

Each diagonal has at least one dot. The dot is intended to signify the lowest value of total cost which can be obtained for a given total average inventory investment. Thus, on the \$40 diagonal the value of $\sum_j T.C_j$ is \$13.10. This value appears at the intersection of average dollar inventory for $P_1 = \$15$ and average dollar inventory for $P_2 = \$25$. Neither value is the optimal value for the individual item considered alone. And so we see, in very specific terms, that the system's optimal does not coincide with the sum of the subsystems' optimals when there is a restriction on the available capital.

A few other examples of the many item effect will be given. We have previously spoken about the difficulty of obtaining the out-of-stock and back-order cost. But the problem is even more difficult than we indicated. Consider the fact that a customer who goes elsewhere to buy a particular item which our company has run out of may never come back to buy that item again. We will assume that he has been buying other items from us as well as the one which we could not supply him with. He may never buy any of those other items from us again. Stated in a more formal way, the inventory policy for each item interacts with that of every other item. Thus a high level of inventory in cigarettes may assure continued demand for newspapers, magazines, stationery, etc. A large inventory investment in a loss leader may be essential to normal demand for other items. Dependencies of this general type also arise, as we have mentioned in Part I, when there is limited storage space for which all items are competing or when combinations of items can be ordered together to obtain discounts or carload shipping rates, and so forth.

59. THE MANY LOCATION EFFECT

Another factor which accounts for the need to consider the whole inventory system as an integrated entity is the existence of many locations, ergo, the *many location effect*. Thus, a company with four or five warehouses ordering individually will lose some of the advantages which are inherent in the operation of a central warehouse.* At the same time, the company will gain other advantages. Some of the disadvantages are additional data-processing costs and higher total investment in inventory. Some advantages are lower shipping costs and greater control. In each case the advantages are compared to the disadvantages for different possible arrangements of the organization of the warehouses. For example, the company could consider having one central warehouse, the remainder being branch warehouses. Another possibility would be two major warehouses each servicing their assigned set of branch warehouses. Accordingly, it can be seen that multiple locations, whether in the form of warehouses or factories, prohibit the company from making an *a priori* decision to consider the total demand for an item as the sum of the demands at each location. It also prohibits an *a priori* decision to treat the demands arising at separate locations as though they were generated by

* See our discussion, pages 162-166.

individual items. The various possibilities should be considered in terms of the whole system. This problem is particularly complicated by the fact that the cost parameters at each location will seldom be the same. Thus, an out-of-stock cost at a minor warehouse with low sales volume will be appreciably less than the out-of-stock cost of a major warehouse with high volume accounts. Prices may be different at each location due to transportation and tax differentials and certainly the cost of ordering is not likely to be the same at each one of the locations.

So once again we are forced to acknowledge that if we do not consider the whole system, but instead install separate inventory plans at each location, we are not likely to optimize the overall company inventory performance.

We could go on at great length showing how the simple inventory approach which treats each item as an entity in itself abuses the principle of optimality. But if we start considering everything even the most patient systems analyst is likely to go berserk. Perhaps eventually we will be able to do this but certainly not at this stage in our evolution of inventory practice. However, that does not mean we have to swing to the other extreme. Interactions and dependencies which appear to be important should be treated in a system's sense. The methods of dynamic programming permit a great deal to be done in this regard. For intensive analysis a computer is essential, since the number of alternatives to be investigated is enormous even though the method of dynamic programming succeeds in reducing the total number of alternatives that would otherwise have to be considered.

60. THE MANY DEPARTMENT EFFECT

Now we shall investigate one further major source of component dependencies. This we shall call the *many department effect*. Here too, because a variety of linkages exist, the inventory system, like the missile system, etc., cannot be successfully reduced to simple, independent subsystems. As we have previously noted, in such systems the optimal design of each component can totally destroy the effectiveness or workability of the integrated system. With respect to many departments the systems concept is essential since the optimal individual design of each department can wreak havoc with overall results. Operations research, as applied to inventory theory, has been especially cognizant of the many department effect and in essence, inventory models treat this aspect of the component interdependencies. Costs are balanced which arise in different departments such as the order cost, storage costs, and outage penalties. If the models did not cut across the organizational boundaries, they would have no value whatsoever.

Let us consider this aspect of the inventory system in more detail. Using Fig. 8-2, we can further explore the question of how the optimum design of individual departments leads to conflict, chaos, and a systems monstrosity.

Figure 8-2 is one possible configuration of interdepartmental flow patterns. With each component we associate certain measures of effectiveness. For example, the sales manager wishes to maximize the quantity-

sales dollars minus sales costs—or he might have as an objective the maximization of sales volume. In any case, his performance is measured by his organization's ability to sell. The performance of the inventory manager is measured in terms of the stock levels he maintains and his avoidance of stock-outs. The production manager desires to minimize his cost record and therefore the measurement of his performance is related to the number of set-ups he uses, his labor costs and the levels of his in-process and raw

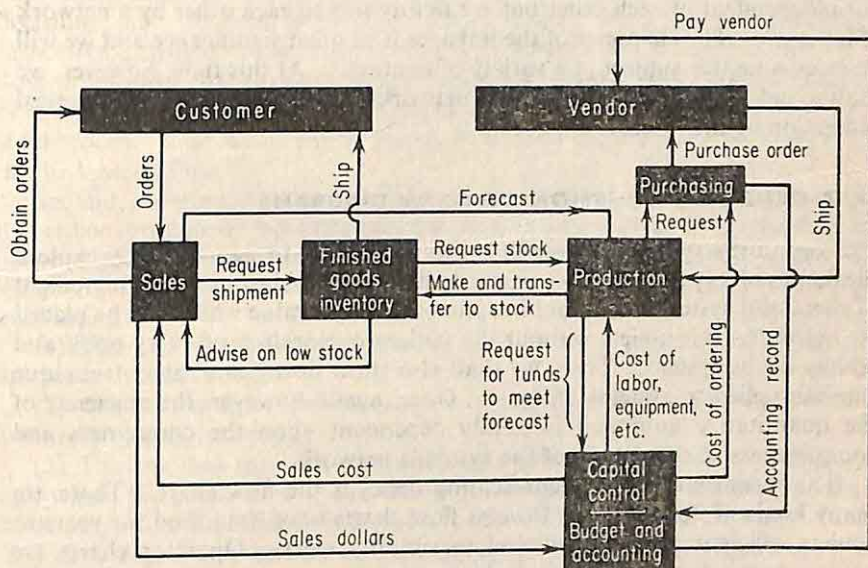


FIG. 8-2 A master block flow diagram.

material's inventories. The allocation of capital is also subject to measures of effectiveness. Here the objective is to maximize the dollars flowing in from sales and minimize the dollars flowing out to production, sales, purchasing, etc. With or without the diagram we see that the objectives of each component are not entirely compatible. We also see that from a systems point of view the measures of effectiveness of the components are badly misused stereotypes. The objectives of the integrated system cannot be satisfied by permitting each component to operate solely in its own interests. The sales manager with a fixed sales force may have to undersell at certain times in order to avoid the costs of back-ordering, expediting, loss of goodwill, costs of overtime, irregular labor force, and so on. Thus, a good measure of effectiveness for his performance must include the effect his actions have on other departments. Similarly, the inventory manager must be prepared to look less than perfect by his own standards. The stock level may have to be higher than he would like at certain times so that set-up costs can be reduced or labor force requirements smoothed. On the other hand,

he may agree with the sales manager when the latter complains that they are too frequently running out of stock because top management refuses to invest more capital in the inventory. But no matter how he feels personally, the well-being of the company must be where his primary loyalty resides. So he accepts the fact that at a convention of inventory managers, he cannot discuss the magnificence of his service record. The same kind of reasoning applies to production, purchasing, finance, etc. We conclude this discussion by stating what should now be obvious if it wasn't before. Only systemic measures of effectiveness are reasonable in those cases where components are not independent of each other but are tightly tied to each other by a network of linkages. The character of the linkages is of great significance and we will be discussing this subject in a variety of contexts. At this time, however, we shall avoid theorizing about system networks and get down to the practical utilization of the systems' approach.

61. A GENERAL DISCUSSION OF FLOW DIAGRAMS

The key to the systems approach is the willingness to painstakingly explore the details of a system. Later on we shall show how ingenuity is an ingredient of successful systems study but in spite of the high value which can be placed on ingenuity, it is useless without the patient exploration of every nook and cranny of the system. Then, we shall also show how quantitative technique can be applied to systems analysis. Once, again, however, the efficiency of the quantitative approach is totally dependent upon the correctness and completeness of the details of the system's network.

The means we use for representing detail is the flow chart. There are many kinds of flow charts. Process flow charts have been used for years to achieve efficient plant layout and to simplify work. Operator charts are used to break down a job into its basic components and sometimes the level of detail which is required is so fine that finger movements are distinguished from wrist movements. Other kinds of flow charts used by industrial engineering are man-machine charts, assembly diagrams, and various kinds of Gantt charts. All of these are familiar production management tools and in their own right they are flow charts. But they do not qualify as legitimate systems analysis unless they comprehend all of the system and not just a component of the system. From our previous discussion of systems we recognize that a system, by definition, must encompass all elements which are interrelated and mutually dependent. Therefore, the construction of a flow chart can provide the necessary level of detail without fulfilling the systems concept. At the same time, systems analysis is hard pressed to proceed without the utilization of flow charts. So flow charts are a necessary but not sufficient requirement of systems analysis.

No matter what form of flow chart is used, the basic and distinguishing characteristic is sequential flow. For the plant layout flow-process chart the sequence may be spatially arranged, while in the operator chart only the time dimension is employed. All flow charts indicate a time-ordered sequence. Figure 8-2 shows time-ordering, but it is too abstract to provide much useful

information. Our purpose in drawing a large-scale flow diagram such as Fig. 8-2 is to indicate how a sweeping view of the system helps to prevent our overlooking components which affect the system to be studied. It is not infrequent that a factor which is ignored, forgotten, or overlooked makes its presence known at a later date in a most unpleasant way.

Let us now note a few particulars about Fig. 8-2. First, the links connecting the various kinds of activities do not all transmit the same kind of thing. Through some of the links pass data, through others pass materials; still others transmit money. A complete flow would show men, machines, decisions, etc. In other words, all of the elements of a system have their respective channels. The flow in any one channel can be affected by the flow in another channel even though the channels transmit different kinds of things. For example, the sales department sends a shipping order to the stock room. The stock room responds to this information flow with a finished goods flow.

Second, the transmission in each channel is seldom continuous. Usually it is either sporadic or regularly repeated at intervals. Note the forecast link between production and sales. This may operate only once each quarter of the year. On the other hand, when production requests raw materials from purchasing:

- (1) The link transmits sporadically if a reorder point (fixed quantity inventory system) is used.

- (2) The link transmits periodically on a fixed cycle if a fixed review period inventory system is used.

- (3) The link may transmit continuously in a continuous process industry; e.g., where the production is that of an oil refinery and the demand for the crude petroleum is satisfied by a pipeline connection to the oil fields.

Third, Fig. 8-2 is one possible mapping of a particular organization's linkages. It is not necessarily a good arrangement. It could be a map of the present situation or, equally, a map for a proposed system. While patience is required to detail the present system, ingenuity is required to rearrange the components, change the connecting links, change the frequency and quality of transmissions, add new components, etc., so as to achieve improved operations with respect to the management's selected measures of effectiveness. Among the many objectives which might exist, the removal of duplication and the reduction of critical delays are familiar system's analysis goals.

Figure 8-2 is a highly compact representation. For example, the storage of inventory obtained from vendors does not even appear. Therefore, it can be assumed to exist within the production box. If the vendors' shipping link had been connected to purchasing, then it would be assumed that the purchasing function subsumed the storage of inventory. An additional link would be required to transfer materials from purchasing to production and all other places requiring such service. In Fig. 8-2, a great number of transactions are obscured by the generalized components and links. We have called this a master block flow diagram because complex components and multiple flows are not detailed (*ergo*, master) and are pictured simply as boxes (*ergo*, block). In other words, this kind of chart emphasizes scope rather

than depth. There is no specific criterion with respect to the amount of detail which should be shown in any flow diagram. It must only be sufficient for its intended purpose. When the intention is to utilize programming and coding for a computer, it must be very detailed. The master and its detail charts, together, form a package which can be used effectively to study the inventory problem.

62. A GENERAL DISCUSSION OF INPUT-OUTPUT MODELS

The inventory system study is used for two different purposes and we do not want to confuse them. The two purposes are:

- (1) for choosing the best possible decision rules for the system as a whole;
- (2) for increasing the processing efficiency of the system as a whole, once an appropriate and satisfactory set of rules have been selected.

The second point will be discussed in the next chapter—our present concern is with the first point.

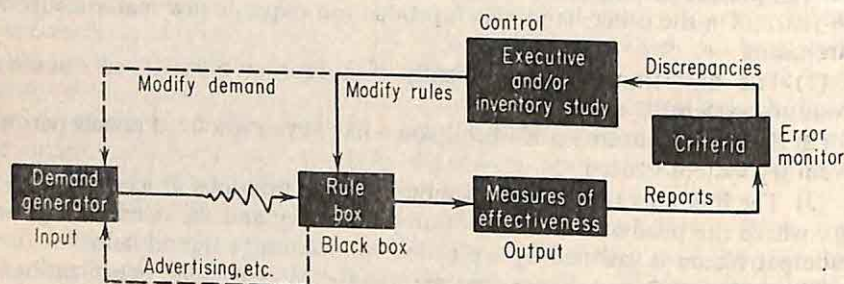


FIG. 8-3 Inventory feedback system.

A generalized version of the kind of system which we are interested in is shown in Fig. 8-3. Let us carefully examine each of the components. To begin with we have a demand generator which is the input to our system. The demand generator is the analog of the marketplace and the purchase pattern over time is the signal which is generated. The reason we talk about a demand generator instead of the marketplace which it is intended to represent, is that when we have designed a device which can imitate or simulate the market pattern, we can run the equivalent of laboratory experiments. Information from the actual marketplace comes to us in real time. That is, it takes a year to get a year's data. With the demand signal generator we can effectively collapse time and produce 50 years of data in a few minutes.

The demand is fed into a rule box. The rule box contains the circuitry which is completely analogous to the company's inventory policy. Each item has its own rule box and of course its own input demand pattern. Another way of stating this, is that the rule box and the inventory models used are equivalent.

From the rule box we obtain the output which consists of several components. Each component is a different measure of the effectiveness of the rule box for the given demand signal. With respect to the inventory problem, the measures of effectiveness would include the service level attained, the number of orders which were required, the average investment in inventory, and the cost of running the system. This last point is the processing and control cost which, as previously explained, is treated in Chapter 9. Figure 8-3 shows that reports describing the measures of effectiveness are issued by the rule box. Accordingly, the measures of effectiveness which are to be used are part of the inventory policy.

An error sensor monitors the measures of effectiveness and reports on all discrepancies between the actual results and the criteria or standards which are the management's objectives. When testing two systems we have in essence two different rule boxes. The same demand pattern is fed into each one and the respective measures of effectiveness are reported for each. Let us say that one is the present system and the second a proposed system. The error sensor, or monitor, will compare the two systems using the output of one (generally the present system) as the criteria or standard against which the second system can be evaluated. If the error detected is favorable to the proposed system, then the proposed rule box will be substituted for the present rule box. Perhaps it is not a sufficient improvement or on the contrary, it might turn out that the proposed system is an impairment. In that case, the control box should call for further study and redesign.

If the system is an operating one and is not being tested then, when the measures of effectiveness meet the criteria, a zero signal (i.e., no signal) is emitted and the control is not alerted to action. The system functions continuously in this fashion until some change occurs or until some set of low probability conditions arises. Then, the error monitor emits a signal which can be likened to an emergency call. The inventory manager is expected to quickly look into the matter and determine what it is that should be done. With a system of this type the executive is free to pursue his executive duties from which basic inventory rules can then be omitted. The manager is called only when sufficiently serious discrepancies occur. This kind of system is frequently called management by exception. It is in effect, a regulator-monitored system* which carries on by itself, utilizing management's prefabricated and repeatable policies and decision rules, until the system encounters factors which it cannot handle. Obviously, at the beginning of inventory systematization the inventory study group is also in the control box. Based on what is learned about the measures of effectiveness, changes are made in the decision or rule box. These changes continue until a satisfactory rule box is designed.

The closed loop,

Rule box → Measures of effectiveness → Criteria

→ Control box → Rule box

is called a feedback loop. Feedback loops are of extreme importance in any

* A regulator is more carefully defined in Chapter 9.

systems analysis. We will become very familiar with them as our discussions continue. Frequently the kind of signal which is fed back can be used to control the output of the system. For example, the guidance system of a plane continually reports discrepancies between the appropriate measures of effectiveness (altitude and heading) and the standard which is the guide beam. The pilot, who is the controller or regulator, directs the plane so as to reduce the reported discrepancies. If an automatic control device is used, then both the rule box and the control box function automatically. That is, a control rule is built into the control box. Inventory systems of this kind have been proposed and we will examine their behavior at a later time. Even in a totally automated system, some device monitors the relevant measures of effectiveness and treats serious discrepancies as emergency situations which require human intervention. Many kinds of loop designs can be used which integrate the executive into the decision process. Some require more executive participation than others. An appropriate design uses the best capabilities of the man vis-a-vis the automated decision-making system. For example, a choice might arise when the measure of effectiveness of service indicates that back orders are occurring with greater frequency than the standard allows. An automated expediting procedure could be called into play. On the other hand, the discrepancy between the output and the criterion might be fed to the automated expeditor until a certain level of discrepancy is reached. At this point the deterioration would be considered to be severe enough to call upon the executive. The executive, in turn, might do nothing more than to invoke the automatic expediting routine. But he might also use special procedures designed specifically to ameliorate the emergency and he might call for a restudy of the decision rules which led to the development of this situation.

In Fig. 8-3, two different dotted lines are connected to the demand generator. One line comes from the rule box and the second from the control box. Now, we know that some controls (albeit, generally weak ones) exist over demand. If one of our criteria is to have a stable work force and it is known that the demand generator produces a seasonal fluctuation, then it is possible to prevent large discrepancies between the criterion and the measure of effectiveness by increasing advertising at the right time. If the situation can be formally anticipated, then it is possible to devise rules which would automatically allocate larger and smaller sums of money for advertising promotion, etc. If not, then the executive might be called upon from time to time to help stabilize the work force by actions calculated to influence the demand generator. Both dotted lines are types of feedback links which are part of feedback loops.

Some feedback loops exist which are not shown in the diagram. The fact that they are hidden within the boxes can be attributed to the low level of detail of the flow diagram. We know when we require the analysis of such loops that we can bring them out of obscurity only if we know how they work. If we want to test different decision rules they must be explicit, detailed, and self-consistent.

These requirements are the same as those of the electrical engineering

field which spawned the input-output model. Electrical engineers call the rule box the "black box." This is because the details of the black box circuitry are either obscured or else they remain to be specified. That is our situation too. Using systems analysis, the relationship between the input and the output can be explored and a set of control objectives can be compared to the actual outputs. On the other hand, the outputs obtained from a specified rule box can be evaluated without knowing anything at all about the actual form of the inputs. Further, adequate decision rule and control boxes can be designed to permit the achievement of objectives.

63. TESTING ALTERNATIVE SYSTEM DESIGNS BY SIMULATION

Let us, without further delay, design a demand generator. We will examine two different types, the first one being the simplest. Assume that a company knows with certainty its monthly demand for an item over the period of one year. This is a series of 12 numbers. Then, let the demand generator repeat this pattern in 12 unit cycles. This data is fed sequentially to the decision rule box. Each time a new number is fed into the system the output is recorded. This generator is merely supplying what is commonly known as an historical forecast.* Let us begin with the simplest possible case. Suppose that the company handles only one item at one location. Suppose further that the present decision rule is: order once each month to cover the next months' requirements. If the demand generator precisely duplicates the actual demand, then the following table describes the result.

TABLE OF MONTHLY ORDER CYCLE SIMULATION
WITH KNOWN FIXED DEMAND SERIES

MONTH	RULE BOX: ORDER x	INPUT: ISSUES	OUTPUTS					
	50		A	B	C	D	E	F
1	35	50	25.0	25.00	1	1.00	4	60.00
2	45	35	17.5	21.25	2	1.70	8	58.20
3	45	45	22.5	21.67	3	2.60	12	58.40
4	50	45	22.5	21.88	4	3.50	16	58.50
5	35	50	25.0	22.50	5	4.50	20	58.80
6	30	35	17.5	21.67	6	5.20	24	58.40
7	40	30	15.0	20.71	7	5.80	28	57.96
8	30	40	20.0	20.63	8	6.60	32	57.90
9	30	30	15.0	20.00	9	7.20	36	56.28
10	40	30	15.0	19.50	10	7.80	40	57.36
11	40	40	20.0	19.55	11	8.60	44	57.36
12	50	40	20.0	19.58	12	9.40	48	57.40
13	35	50	25.0	20.00	13	10.40	52	57.60

What we have just done is known as simulation. The general meaning

* The historical forecast takes account of seasonality but no allowance is made for trend. See relevant discussion in Chapter 9.

of the word is as good a definition as any. Our dictionary states that "to simulate" is "to assume the appearance of, without the reality." No one could confuse our rule box and demand generator with reality. They are an attempt to assume the appearance of reality. In this way we can observe the performance of an analog version of the real thing. And the advantage is that we can compress time and observe results that would otherwise require prohibitive periods of time. Further, we can experiment without actually putting our head into the lion's jaws. No executive likes to tamper with the company's policies, particularly when such tampering might invite catastrophe. But the simulation procedures of systems analysis permit long shots to be tested and encourage a sound experimental program.

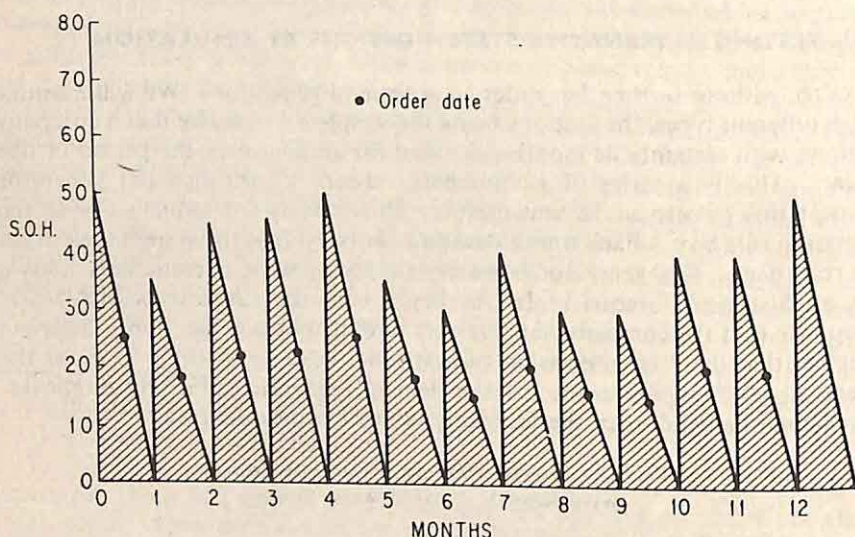


FIG. 8-4 Simulation result with monthly order cycle and known fixed demand series.

Now, let us look at the outputs obtained from our first simulation. The output columns are designated by letters which should be translated as follows:

- A: Average inventory in units,
- B: Running average inventory in units,
- C: Cumulative number of orders ($L.T. = \frac{1}{2}$ month),
- D: Running carrying cost of average dollar inventory ($c = \$2.00$ per unit, $C_c = .24$ per year),
- E: Cumulative order cost ($C_r = \$4.00$ per order),
- F: Total cost (sum of $D + E$) adjusted to yearly basis.

An executive in the control box would study the outputs, that is, the various measures of effectiveness (henceforth called *moe*). Consulting a visual interpretation of the results, as in Fig. 8-4, makes it somewhat easier to

understand what has happened. The executive would undoubtedly be very satisfied with the status of the average inventory. He would have no complaint about the stock-outs since none occur. But he should be dissatisfied with the size of the ordering costs. He observes that the carrying cost tends to be much smaller than the ordering cost, which indicates to him that he might be doing a better job. Further, the carrying cost moves closer to the ordering cost as the order size increases. Believing that a lower total cost could be obtained, he calls for study of the situation and he receives the following advice: set the order quantity by means of the economic lot size formula:

$$x_o = \sqrt{\frac{2(470)(4)}{2(.24)}} = \sqrt{7833} = 88 \text{ units}$$

where the demand per year equals 470 units and the remaining parameters are the same ones which were used before.

Then the number of orders per year will be:

$$n = \frac{z}{x} = \frac{470}{88} = 5.3 \text{ orders per year}$$

$$t = \frac{12}{n} = 2.25 \text{ months between orders}$$

Furthermore, a reserve stock will be carried. Previously, the demand was assumed to be known with certainty. The demand generator merely did what it was told to do. But the inventory analyst is skeptical. He believes that precautions must be taken since demand in the future will vary and will not exactly duplicate the previous 12 month series. Consequently he mixes up the order in which demand was experienced and obtains the following distribution.

ISSUES PER MONTH	FREQUENCY	PROBABILITY	CUMULATIVE PROBABILITY
30	3	.250	.250
35	2	.167	.417
40	3	.250	.667
45	2	.167	.834
50	2	.167	1.001
	12	1.001	

We are intentionally using a distribution with an irregular shape. In practice such distributions are frequently encountered. Various assumptions can be made to smooth them or place them in a family of known characteristics. For this distribution then, the expected number of issues per month is found to be:

$$\bar{z}_m = 39.2$$

The standard deviation of monthly demand is:

$$\sigma_m = 7$$

Lead time is constant at one-half month. Therefore, the expected demand in the lead time period is:

$$\bar{z}_{L.T.} = \frac{39.2}{2} = 19.6$$

The standard deviation of demand in the lead time period:

$$\sigma_{L.T.} = \sqrt{\frac{1}{2}} \sigma_m = \frac{7}{1.4} = 5$$

A service level of about .05 is desired; i.e., $\alpha = .05$. If we assume that the Poisson family approximates our distribution, then we can estimate that a

Order date

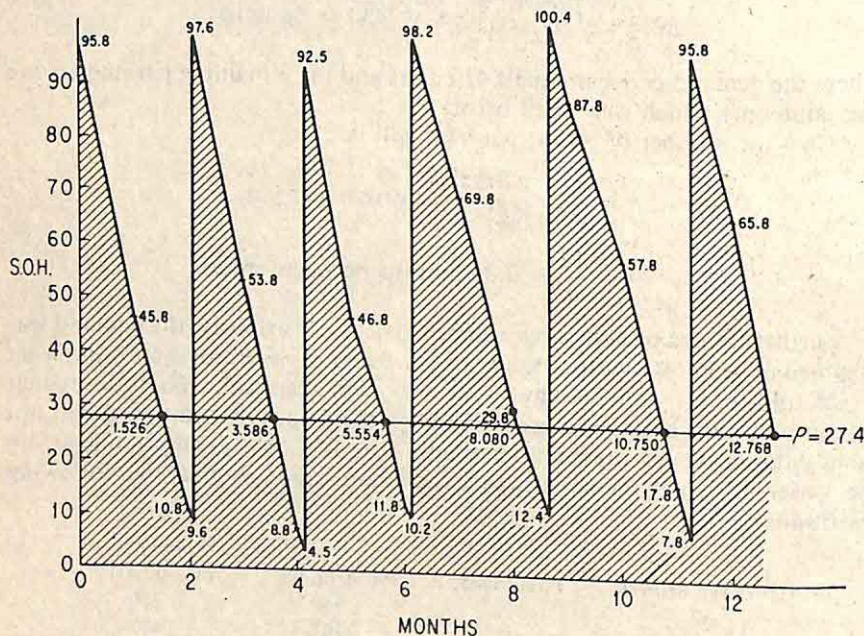


FIG. 8-5 Q-model simulation with fixed demand series.

value of 27.4 will cut off the tail of the distribution leaving a probability of .05 of exceeding that value. Consequently, the reserve stock corresponding to $\alpha = .05$ will be $27.4 - 19.6 = 7.8$ units ($K = 1.56$ units).

Our rule box is completed and can be summarized:

- (1) Order 88 units when the reorder point is reached.
- (2) The reorder point will be reached approximately 5.3 times per year.
- (3) The reorder point,

$$P = \bar{z}_{L.T.} + \text{Reserve stock}$$

and

$$P = 19.6 + 7.8 = 27.4 \text{ units}$$

We will now run through the simulation of this model which we recognize to be a Q -type model as described in this text. We use the fixed demand pattern of the historical forecast. The results are shown in Fig. 8-5.

It is usually helpful to have a simulation diagram accompany the sequence of computations, but it is not a necessity since the output measures, which are the results we require, must still be calculated in order to evaluate the model. The steps to derive the table follow directly from the definition of the rule box. The moe are computed from the accompanying table of the Q -model simulation.

TABLE OF Q -MODEL SIMULATION
WITH FIXED DEMAND SERIES

MONTH	S.O.H.	ISSUES	ORDERS	RECEIPT
0	95.8			
1	45.8	50.0 50		
*1.526	27.4	18.4}	88	
2	10.8	16.6}		
2.026	9.6-97.6	1.2}		88
3	53.8	43.8}		
*3.586	27.4	26.4}	88	
4	8.8	18.6}		
4.086	4.5-92.5	4.3}		88
5	46.8	45.7}		
*5.554	27.4	19.4}	88	
6	11.8	15.6}		
6.054	10.2-98.2	1.6}		88
7	69.8	28.4}		
8	29.8	40.0 40		
*8.080	27.4	2.4}	88	
8.580	12.4-100.4	15.0}		88
9	87.8	12.6}		
10	57.8	30.0 30		
*10.750	27.4	30.4}	88	
11	17.8	9.6}		
11.250	7.8-95.8	10.0}		88
12	65.8	30.0}		
*12.768	27.4	38.4}	88	
13	15.8	11.6}		
13.268	6.4-94.4	9.4}		88
		25.6}		

S.O.H. = stock on hand.

* = reorder point

- (1) Begin with stock on hand (S.O.H.).
- (2) Subtract issues of the first month.

(3) If the reorder point is not reached, subtract next month's demand. Continue subtracting each month's demand until the reorder point is either reached or exceeded.

(4) When the reorder point is reached, place an order which will definitely be received .5 months later.

(5) Adjust stock on hand when delivery is made and continue to withdraw items from stock at the specified monthly demand rates.

What are the results of this simulation? We will compute only the end of year moe and none of the interim values. Further, we will not bother to report moe *A* since it is implicit in the derivation of moe *B*.

B: Running average inventory in units = 54.5.

C: Cumulative number of orders = 5.6.

D: Running carrying cost of average dollar inventory = \$26.16.

E: Cumulative order cost = \$22.40.

F: Total cost = \$48.56.

The number of orders is quickly counted either from the diagram or the table. The fractional amount expresses the fact that the last order of the year does not occur on the last day of the year. The cumulative average inventory is obtained as follows:

$$\begin{array}{rcl}
 (95.8 + 45.8)/2 \times 1.000 & = & 70.8 \\
 (45.8 + 10.8)/2 \times 1.000 & = & 28.3 \\
 (10.8 + 9.6)/2 \times .026 & = & .3 \\
 (97.6 + 53.8)/2 \times .974 & = & 73.7 \\
 (53.8 + 8.8)/2 \times 1.000 & = & 31.3 \\
 \vdots & & \vdots \\
 \text{etc.} & \Sigma = & 12.000 \quad 654.0 \\
 & & \vdots \\
 & & \frac{654}{12} = 54.5
 \end{array}$$

D, *E*, and *F* are computed in the usual way once we have been supplied with *B* and *C*.

Apparently, the advice given to the executive was good. The total cost for the year has been reduced from \$57.40 to \$48.56, which is a 15.4% improvement. The improvement was obtained by decreasing the number of orders from 48 to 5.6. This is an 88% decrease both in the number of orders and in the ordering cost. At the same time the average number of units in inventory increased from 19.58 to 54.50. The carrying cost of this inventory increased from \$9.40 to \$26.16, which is about a 175% increase. The original policy was quite far off the mark. With the new plan, as was the case with the old one, no stock-outs occurred. The reserve stock level was broached once and just reached one other time.

Let us stop for a moment and compare these two models and their respective rule boxes and demand patterns. Our first model could only be realistic if the company has contracts which permit the exact estimation of monthly requirements. But obviously, by following a monthly ordering cycle, the company is not using its prize possession—certainty of demand—to

its own best advantage. The company does not even do as well as it could have, had it thrown away the time-sequential information and proceeded as though it was operating under risk conditions. We know from our previous discussion in the text that a dynamic program could be used to produce the best possible ordering policy.* So this additional alternative presents itself. The dynamic program, if applicable to many different items, would be quite onerous for manual computation but could be efficiently performed on a computer.

64. MONTE CARLO SIMULATION—THE Q-MODEL

The second kind of demand generator that we will discuss is more elaborate than the fixed series type. It also does away with the contract restriction which we specified in the previous section. We assume that the company has the same one year of data to work with. Ordinarily, no one would be particularly happy with such a small sample but, on the other hand, sometimes even less than a year's supply of data are all that is available.

The probability distribution of monthly demand obtained in the previous section is shown in the first three columns below:

ISSUES PER MONTH	PROBABILITY	CUMULATIVE PROBABILITY	MONTE CARLO NUMBER
30	.250	.250	000-249
35	.167	.417	250-416
40	.250	.667	417-666
45	.167	.834	667-833
50	.166	1.000†	834-999

We hypothesize that no sequential demand dependencies exist. In other words, each month's demand is independent of the demand that occurred in the preceding month. Then we want our demand generator to produce a random series of monthly demands which we can use as input to the various rule boxes that we are testing. If we simulate enough sequences and observe the moe which are obtained, we can then evaluate the various alternatives with confidence that each system is closely described by the results which we have obtained from the simulation.

The procedure of generating a demand series from a demand distribution requires the assignment of Monte Carlo numbers. This is the last column in the table above. It will be noted that exactly the right proportion of different numbers has been assigned to each issue level. Thus, 250 different, three-digit numbers are used to represent the expected frequency of occurrence of 30 issues per month. It does not matter which numbers are used in each

* The use of the dynamic programming method produces a substantial reduction in the T.C. We suggest that the reader perform the necessary computations as an exercise.

† We have adjusted the probability of 50 issues per month in order to obtain unity summation. We can always carry more decimal places if greater accuracy is required. In this case, the added precision is not warranted.

category as long as no number is ever used more than once and as long as each category is correctly weighted. The assignment of Monte Carlo numbers in the table above meets this description. The (issues = 35) category is weighted with 167 Monte Carlo numbers out of the total of 1000 which cover the whole distribution.

If we find a way to generate three-digit numbers so that any number is equally likely, we can then take this random number and match it against the assigned Monte Carlo numbers. If such a random number was 643, we then read directly from the table that the Monte Carlo number 643 is equivalent to 40 issues per month. Since any number, 000-999, is equally likely to appear, then a category containing 250 Monte Carlo numbers is expected to be chosen five times more frequently than a category which contains 50 Monte Carlo numbers. In this simple way we can continue to take samples from the distribution.

But we still must find a way to obtain the random numbers. There are various methods for producing sufficiently random numbers. With the assistance of a computer, it is possible to spew forth great quantities of satisfactory random numbers. Many times, when random numbers are required in a simulation, the computer is programmed to produce such random numbers as they are needed. An alternative procedure is to obtain a set of random numbers and then store them in the memory unit of the computer. This is like a table of random numbers. The disadvantage of using such a stored table in a computer is that it uses up memory capacity which may be required for other parts of the program. However, a table of random numbers is by far the simplest approach for manual computation. We present below an illustrative set of such random numbers.*

470	881	121	684	388	833	845	472	647	728
129	553	402	157	902	618	870	098	099	588
116	610	118	022	636	726	493	167	879	065
293	534	399	781	511	765	222	263	465	896

This table can be directly translated into the sample of usage rates, namely,

40	50	30	45	35	45	50	40	40	45
30	40	35	30	50	40	50	30	30	40
30	40	30	30	40	45	40	30	50	30
35	40	35	45	40	45	30	35	40	50

We group these 40 sample results and obtain their distribution.

ISSUES PER MONTH	PROBABILITY
30	.275
35	.125
40	.300
45	.150
50	.150
	1.000

* A table of random numbers will be found on page 337.

Even with the small sample that we have used, the derived distribution is quite similar to the original distribution. And we now have the advantage of a sequence of demands which are independent of each other. We will use the first 12 randomly generated demands as the input to our two previous rule boxes.

For the monthly ordering policy rules we obtain:

TABLE OF MONTHLY ORDER CYCLE SIMULATION
WITH RANDOM DEMAND SERIES

MONTH	S.O.H.	ISSUES	ORDER	RECEIPT
0	50-10	40	35	..
1	10-45	50	45	35
2	(-5)-40	30	45	45
3	10-55	45	50	45
4	10-60	35	35	50
5	25-60	45	30	35
6	15-45	50	40	30
7	(-5)-35	40	30	40
8	(-5)-25	40	30	30
9	(-15)-15	45	40	30
10	(-30)-10	30	40	40
11	(-20)-20	40	50	40
12	(-20)-30	..	..	50

Figure 8-6 presents the sad picture, but to precisely evaluate how bad it

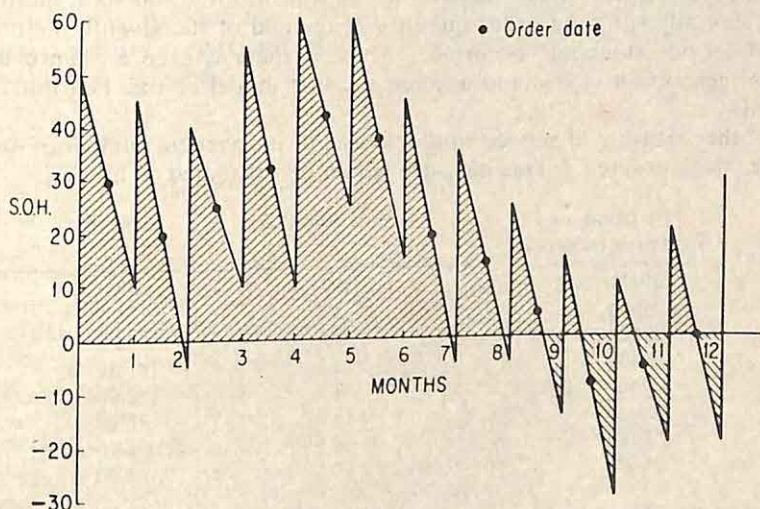


FIG. 8-6 Simulation result with monthly order cycle and randomly generated demand series.

really is, we must derive our moe. (The items in the stock-on-hand column which bear minus signs and are enclosed in parentheses are stock-outs.) Our

computations for all the moe except service level percent are not shown since they require no different procedures from those which were previously employed.

There are several different measures which can be used to describe the kind of service rendered. For example, we could use the number of months in the year that any kind of stock-out occurs. In our case, this would be 7 out of 12 or 58%. It will be recognized that this is equivalent to α since an outage resulted in 7 out of 12 lead time periods. This is more than ten times the α which management had specified. Another possible service moe is the percent of time that stock-outs exist. This is:

$$\frac{5}{50} + \frac{5}{50} + \frac{5}{40} + \frac{15}{40} + \frac{30}{45} + \frac{20}{30} + \frac{20}{40} = 2.5$$

and

$$\frac{2.5}{12} = .21$$

We obtain in this way further confirmation of the fact that a monthly ordering policy which is preset on the basis of an historical forecast is a bad policy indeed—when, in fact, the demand is randomly generated. Of course, when the executive sees what is happening (*if he sees what is happening*), he would certainly modify the historical forecast. Yet if he had adjusted the forecast at the end of the second month, he would have only helped to increase the high inventory levels which occurred in the fourth, fifth, and sixth months. But by not adjusting the order quantity at the end of the seventh month—a rash of serious stockouts occurred. Thus, if there is even a chance that random generation of demand applies, the first model or rule box must be rejected.

Another measure of service would consist of the average number of items that are back-ordered. This measure would be calculated as follows:

INTERVAL OF SHORTAGE (MONTHS)	AVERAGE SHORTAGE	PRODUCT
19/2	0.0	0.00
5/50	2.5	.25
5/50	2.5	.25
5/40	2.5	.31
15/40	7.5	2.80
30/45	15.0	10.00
20/30	10.0	6.70
20/40	10.0	5.00
12		25.31

and, $25.31/12 = 2.11$ average number of units back-ordered at any time. We can also find the average number of units back-ordered assuming an outage exists. This is:

$$\frac{25.31}{2.5} = 10.15 \text{ units}$$

If all demands are not made in unit quantities, then it would be possible to determine the percent of total demands that cannot be filled immediately which is another possible variant of a service moe.

A more precise statement of the outage measures given above will shed additional light on the meaning and dimensions of such measures. Consider the outages which occur at the end of the second month. We have:

NUMBER OF UNITS SHORT	LENGTH OF TIME (MONTHS)	UNIT-MONTHS SHORT
1	1/50	1/50
2	1/50	2/50
3	1/50	3/50
4	1/50	4/50
5	1/50	5/50
		15/50

We will have occasion to use this kind of measure again in an unexpected context so let us be certain as to its meaning. If a shortage of 2 units persisted over the period of one year, it would result in a service measure of $2 \times 12 = 24$ stock-out unit months. The penalties paid for shortages would, in many cases, increase in proportion to the stock-out unit months rather than in proportion to any of the previous measures we have discussed. For our example, the total stock-out unit months is computed to be 26.56. This figure is more precise but comparable to the 25.31 stock-out unit months previously obtained.

End of month number	2	7	8	9	10	11	12	TOTAL
Stock-out unit months	15/50	15/50	15/40	120/40	465/45	210/30	210/40	26.56

This is equivalent to a constant shortage of 2.21 units over the course of an entire year. Viewed in this way the stock-out situation does not seem so serious. Clearly, the correct outage measure is the one which reflects the basis upon which the outage penalty is levied. Our first measure indicated a very serious situation—i.e., 58% of all lead times experience at least one outage. Our second measure, outages exist 21% of the time, is certainly not a good showing. It reflects the fact that at least $\frac{1}{5}$ of the time some customer is unhappy. Our third measure, an approximate constant shortage of two units does not strike us as forbidding. But if many irate customers share in the creation of the two-unit outage, then this small number may mask great penalties.

How about the other moe? There are 12 orders, as before, costing \$48. The average inventory is 20 units with an average yearly dollar carrying cost of \$9.60. This is just about the same investment cost which was experienced when the actual sequence of demand was able to be predicted.

At this point, the advantages of the simulation procedure have become

much clearer. We will take our analysis a step further by applying the random demand generator as an input to our Q -rule box. Figure 8-7 pictures the results we obtain this time.

TABLE OF Q -MODEL SIMULATION
WITH RANDOM DEMAND SERIES

MONTH	S.O.H.	ISSUES	ORDER	RECEIPTS
0	95.8			
1	55.8	40.0 40		
*1.509	27.4	28.4	88	
2	5.8	21.6		
2.009	5.5-93.5	.3		88
3	63.8	29.7		
*3.810	27.4	36.4	88	
4	18.8	8.6		
4.310	7.9-95.9	10.9		88
5	71.8	24.1		
*5.987	27.4	44.4	88	
6	26.8	.6		
6.487	2.4-90.4	24.4		88
7	64.8	25.6		
*7.935	27.4	37.4	88	
8	24.8	2.6		
8.435	7.4-95.4	17.4		88
9	72.8	22.6		
10	27.8	45.0 45		
*10.013	27.4	.4	88	
10.513	12.8-100.8	14.6		88
11	85.8	15.0		
12	45.8	40.0 40		

The number of orders can be approximated at 5.8. The average number of units carried in the inventory is equal to $639/12 = 53.3$. So we can summarize our results:

B: Running average inventory in units = 53.3.

C: Cumulative number of orders = 5.8.

D: Running carrying cost of average dollar inventory = \$25.58.

E: Cumulative order cost = \$23.20.

F: Total cost = \$48.78.

As we would have expected these results are very similar to those previously obtained for this same Q -model rule box. This model was designed to accept the signal of a random demand generator. Consequently, while there will be variations over time as different series of demands occur, nevertheless, the performance in the long run will average out in the neighborhood of the theoretical moe. The fact that the first series was fixed by a

management decision does not influence the fact that it represents just one possible random sequence which could by chance have been produced by the demand generator. As a matter of fact, unless more than one year's data are available, it is best to assume randomness. Of course, contracts or similar assurances of fixed sequences would nullify the remarks above. In our later

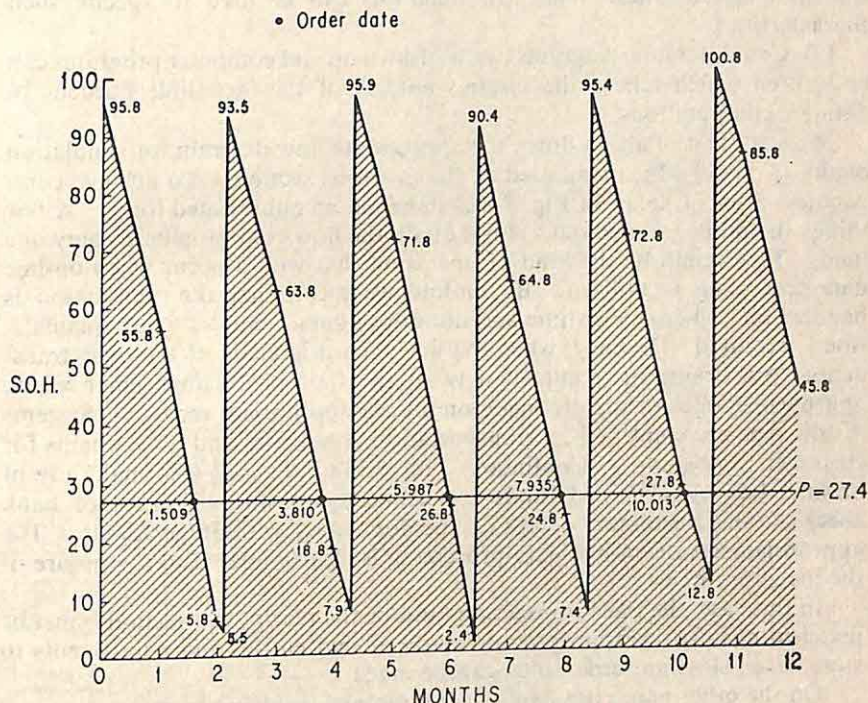


FIG. 8-7 Q-model simulation with random demand series.

discussions of forecasting we will be able to amplify a good deal on these conclusions. Such amplification is dependent upon the development of certain notions which we will not introduce at this time.

One further point deserves mention. No stock-outs occurred in either simulation run although the rule box had been designed to permit an outage one time out of every 20 lead time periods. Since only five orders were placed in each simulation it is not surprising that no stock-out was experienced. This indicates the need for lengthy simulations—easily accomplished on a computer.

65. DETAILED FLOW DIAGRAMS

Each individual demand input requires a new computation to be made. And each set of computations is called an iteration. Detailed flow diagrams

are generally used by systems analysts to represent the iterative procedures used for simulation as well as for other kinds of computations which require that a series of calculations be repeated. The purposes served by such flow diagrams are at least threefold:

(1) Clarification of points that may not even seem to exist is forced upon the analyst.

(2) The effects of feedback loops can be investigated as well as other systems characteristics. The flow diagrams can be used to specify such characteristics.

(3) Computer flow diagrams can be drawn up and computer programs can be written which relieve the systems analyst of the incredible burdens of iterative computations.

We shall first of all set down the appropriate flow diagram for simulation of the Q -model which was used in the previous section. To achieve compactness a lot of detail in Fig. 8-8 is shown in an abbreviated form. A few things should be pointed out. First of all, the flow chart applies to only one item. This would be the kind of operation that would occur when on-line data-processing is utilized. By on-line we mean that the transaction is handled on real company time and not after hours. But a further qualification is required. Namely, when the first record is made of the issue transaction, it is transmitted immediately to the data-processing center where appropriate adjustments are made on all the appropriate records. Systems of this type are employed by transportation companies and hotel chains for customer reservations. Continuous processing industries also make use of on-line data systems. Methods for achieving on-line clearance of bank checks so that customary delays can be done away with are available. The appropriateness of on-line data-processing is clearly related to the nature of the inventory items.

In this case, the issue transaction (and/or the receipt transaction) must be matched against the appropriate records so that entries and adjustments to stock level, stock on order, etc., can be made.

On the other hand, if a batch of transactions applying to many items are saved up,* then a typical data-processing run will involve matching each item against its appropriate record and providing loop sequences which will terminate the operation when all transactions have been processed.† This is the more usual case. But it does an injustice to the Q -model which in theory is a perpetual inventory system requiring on-line, immediate data-procedure.

Figure 8-8 indicates that two tests are made with respect to the stock on hand plus the stock on order. The first test is whether or not the reorder point has been reached or exceeded. The second test is used to determine how badly depleted the stock position is. If a serious shortage appears to be imminent then the routine calls for an expediter.

* See discussion on optimum run sizes in Gregory and Van Horn, *Automatic Data-Processing Systems* (San Francisco: Wadsworth Publishing Co., Inc., 1960), pp. 590-593.

† The appropriate flow diagram will be found in Chapter 9 at which point the emphasis is on processing rather than decision characteristics.

A most important component of our flow diagram is the updating routine which is used. Updating can take place with every transaction. It could take place with every twentieth transaction. Another possibility is that

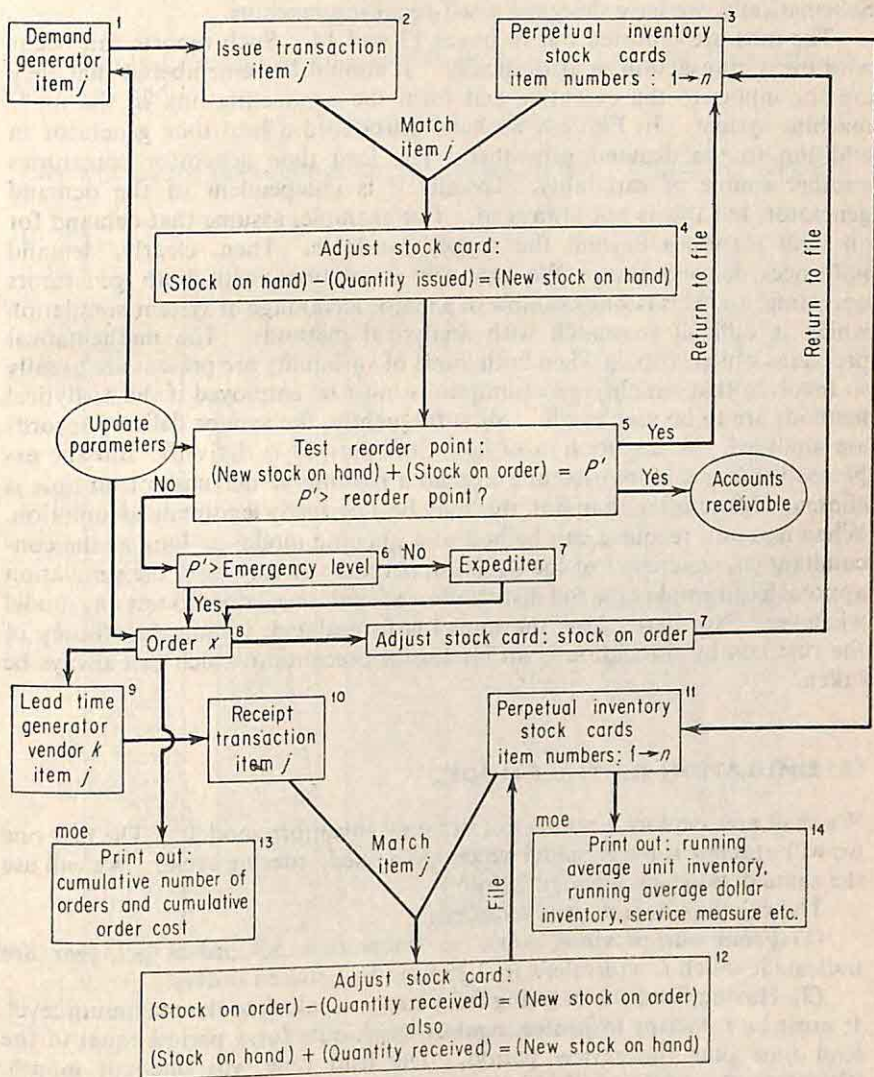


FIG. 8-8 Detailed flow diagram of the Q-model simulation.

whenever the reorder point is reached our rule box calls for updating on the basis of everything which has previously transpired. Still another possibility is to update every day, every week, every month, every six months, and so on.

Presumably, no matter what updating system is chosen it best fits the processing routines used and the stability characteristics of the items involved. In effect, updating is equivalent to monitoring the reorder point, the order quantity, and the reserve stock. Changes in these factors are introduced automatically by updating which in itself can be an automated procedure. Schematically, we have designed a self-repairing machine.

The moe are reported out in boxes 13 and 14. Such reports can occur with every transaction or periodically. It should be remembered that they are the inputs to the executive and form the connecting link of the man-machine system. In Fig. 8-8 we have introduced a lead time generator in addition to the demand generator. The lead time generator constitutes another source of variability. Usually it is independent of the demand generator, but this is not always so. For example, assume that demand for an item increases beyond the supply available. Then, clearly, demand influences delivery time. We can run simulations with both generators operating, and this is one example of a major advantage of system simulation which is difficult to match with analytical methods. The mathematical problems which crop up when both kinds of variability are present are usually so involved that simplifying assumptions must be employed if the analytical methods are to be used at all. Most frequently, the vendor delivery records are analysed. A distribution of lead time periods is derived. But the expected value is seldom used and instead a reasonable maximum lead time is chosen. More often than not, this may be a perfectly legitimate assumption. When it is not, recourse can be had to a queuing model—as long as the concomitant characteristics of the situation permit. In any case, the simulation approach can employ the full distribution of lead time values to test any model whatever. No matter how the model is formulated, testing the efficacy of the rule box by simulation is an intelligent precaution which can always be taken.

66. SIMULATION OF THE P-MODEL

We shall now explore a number of different inventory models. The first one we will simulate is the *P*-model which has a fixed ordering cycle. We will use the same data which we used before.

The rule box definition is required:

(1) From our previous work we know that 5.3 orders per year are indicated, which is equivalent to 2.25 months between orders.

(2) Having fixed the ordering frequency we calculate the maximum level. It must be sufficient to protect against stock-outs for a period equal to the lead time plus the review period. The lead time was one-half month. Therefore, the total period to be covered is $2.25 + .50 = 2.75$ months. The expected demand in the total period will be:

$$\bar{z}_{R.P. + L.T.} = (2.75)\bar{z}_m = (2.75)(39.2) = 107.8 \text{ units}$$

We will use the normal approximation to the Poisson distribution since

the demand is large enough to warrant this assumption. For $\alpha = .05$ we obtain a K factor of 1.64. The standard deviation of the Poisson distribution is equal to the square root of the expected value of the distribution. In our case, this is 10.4 units. Therefore, the reserve stock required to achieve the stated protection level will be:

$$(1.64)(10.4) = 17.1 \text{ units}$$

(3) The last element we must specify for our rule box is M , the maximum stock level, and

$$M = 107.8 + 17.1 \approx 125 \text{ units}$$

We will use the randomly generated demand series. Figure 8-9 illustrates

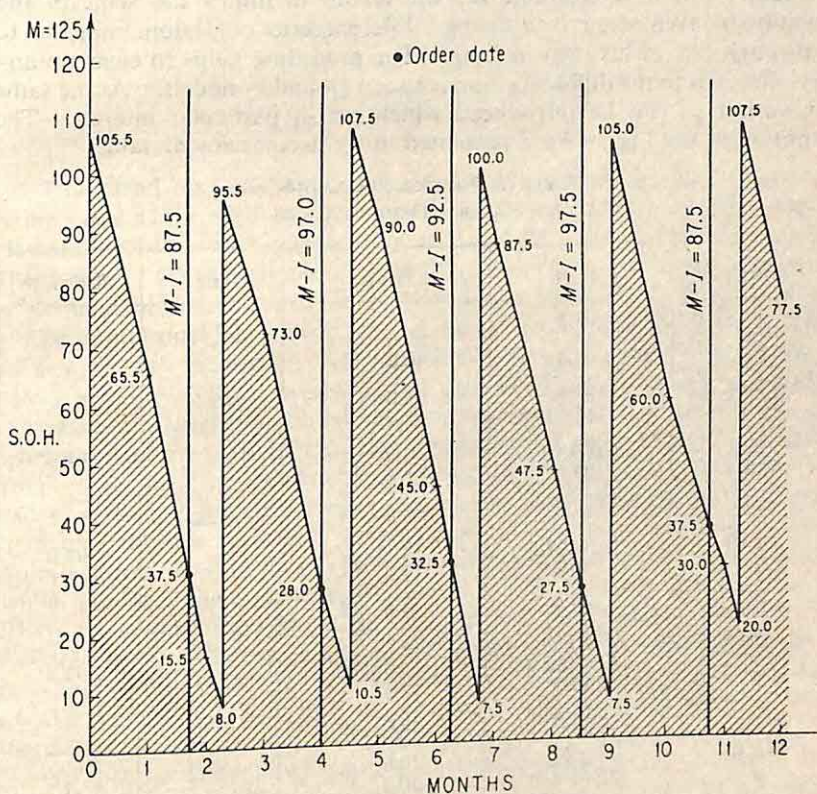


FIG. 8-9 P-model simulation with random demand series.

the result of the simulation. Once, again, no stock-outs have been experienced. There is no question about the number of orders. They were preset and controlled at 5.3 orders per year. To the eye the average inventory seems to have increased slightly. Let us now examine the moe.

B: Running average inventory in units = 61.4.

C: Cumulative number of orders = 5.3.

D: Running carrying cost of average dollar inventory = \$29.47.

E: Cumulative order cost = \$21.20.

F: Total cost = \$50.67.

There has been a small increase in the average inventory, but since the number of orders is held constant, the rise in the total cost will not be severe. Nevertheless, we recognize that the *P*-model has not performed quite as well as the *Q*-model.* The use of the *P*-model would be justified by savings realized from a fixed order period and the abolition of a perpetual inventory system—which is required for the *Q*-model. The simulation procedure with the *P*-model rule box and the random demand generator imposes no new difficulties. But it is desirable for the reader to follow the steps of the simulation or, even better, by utilizing a different series of random numbers, to run through one of his own making. The procedure helps to clear up any uncertainties as to the differentiation between *Q*- and *P*-models. At the same time, variations can be introduced which are of particular interest. The computations for Fig. 8-9 are presented in the accompanying table.

TABLE OF *P*-MODEL SIMULATION
WITH RANDOM DEMAND SERIES

MONTH	S.O.H.	ISSUES	ORDERS	RECEIPT
0	105.5			
1	65.5	40.0 40		
*1.75	37.5	28.0 } 50	87.5	
2	15.5	22.0 }		
2.25	8.0-95.5	7.5 } 30		87.5
3	73.0	22.5 }		
*4	28.0	45.0 45	97.0	
4.50	10.5-107.5	17.5 } 35		97.0
5	90.0	17.5 }		
6	45.0	45.0 45		
*6.25	32.5	12.5 } 50	92.5	
6.75	7.5-100.0	25.0 }		92.5
7	87.5	12.5 }		
8	47.5	40.0 40		
*8.50	27.5	20.0 } 40	97.5	
9	7.5-105.0	20.0 }		97.5
10	60.0	45.0 45		
*10.75	37.5	22.5 } 30	87.5	
11	30.0	7.5 }		
11.25	20.0-107.5	10.0 } 40		87.5
12	77.5	30.0 }		

* See relevant discussion in Chapter 5.

67. SIMULATION OF P-MODEL WITH VARIABLE LEAD TIME

We introduce one new variation in this simulation. Namely, a lead time generator is applied to the rule box. We want to observe the outputs of the system when this new condition is added. Let us assume that lead time is distributed in the following manner:

L.T. (MONTHS)	L.T. PROBABILITY	CUMULATIVE L.T. PROBABILITY	MONTE CARLO NUMBERS
0	.00	.00	..
1/4	.23	.23	00-22
1/2	.60	.83	23-82
3/4	.10	.93	83-92
1	.07	1.00	93-99
	1.00		

Thus, lead time has an *expected value* of one-half month. This is the same value as the constant lead time which we used before. The standard deviation of this distribution, $\sigma_{L.T.} = \frac{1}{2}$ month. Two alternatives present themselves. On the one hand, we could attempt to combine the distributions of demand and lead time, but this is not always possible. When it is possible the computational burden is large and the results not always useful. Instead of dealing with the product function of these two random variables, we can select a value, $L.T._{max}$, which would be $.50 + K\sigma_{L.T.}$ * This succeeds in increasing the reserve stock and preventing stock outages which might arise when the lead time—by chance—takes on one of its larger values. The adverse effect is obvious. If we choose too large a value for $L.T._{max}$, we then incur excessive carrying costs. Differentiation between practical and theoretical systems is always in order. Overanalysis of the L.T. distribution can be pointless in view of the variety of other uncertainties and the lack of precision which pervades the inventory decision field. Simulation offers a relatively inexpensive substitute for large investments in analytical studies. For example, we will choose $L.T._{max} = \frac{3}{4}$ month. Our simulation produces the relevant output moe. If we think we can better these results we select another value for $L.T._{max}$ and run another simulation. Proceeding in this manner, it is possible to approach an optimal solution.

We generate the following random number series:

90 26 18 87 00 98 09 95 88

These correspond to lead times of:

$\frac{3}{4}$ $\frac{1}{2}$ $\frac{1}{4}$ $\frac{3}{4}$ $\frac{1}{4}$ 1 $\frac{1}{4}$ 1 $\frac{3}{4}$

It is immediately obvious that the sequence contains low probability

* The significance of K with respect to protection is lost however since neither the demand reserve stock K nor the lead time reserve stock K identify the cutoff value of the multiple distribution of demand and lead time.

events. Although it is rare, it is nevertheless a situation which can arise and must be expected. Reference to Fig. 8-10 shows what occurs when the P -model experiences such a run of lead times that deviate considerably from

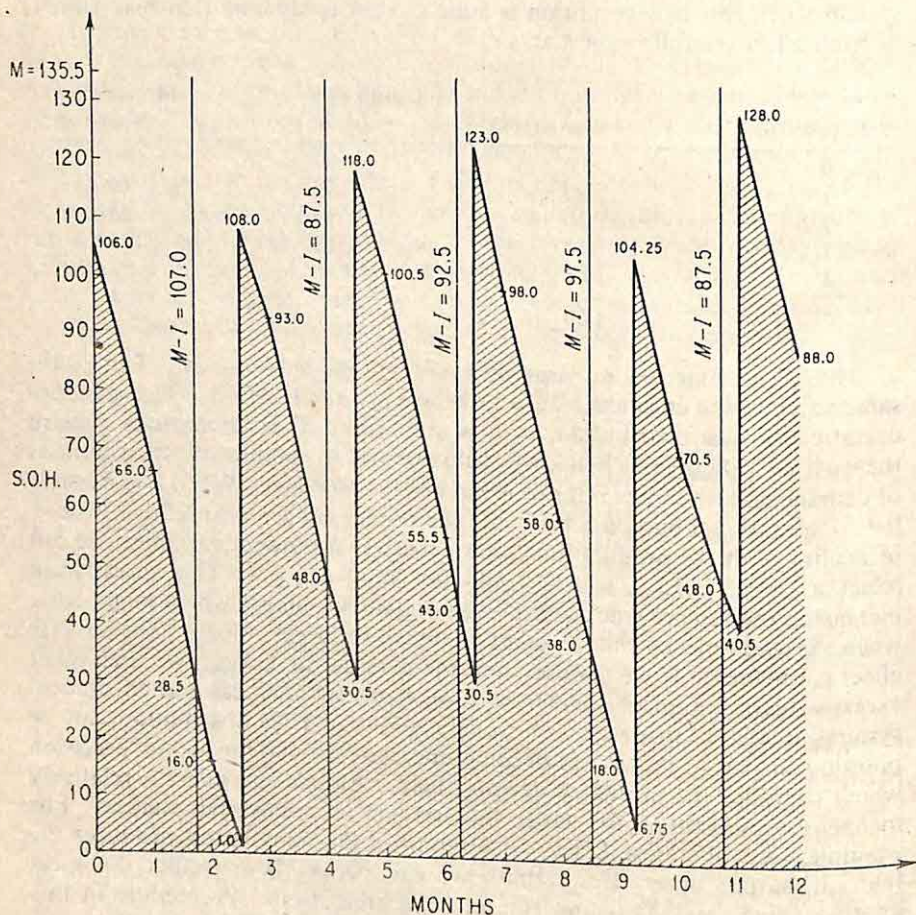


FIG. 8-10 P -model simulation with random demand series and random lead time series.

the expected value. Remember, we have made an additional safety allowance by obtaining the demand in a longer lead time period (specifically $L.T_{\max}$). Consequently, our rule box is:

- (1) $n = 5.3$ orders per year (unchanged).
- (2) $t = 12/n = 2.25$ months between orders (unchanged).
- (3) $M = \bar{z}_{R.P. + L.T.} + K\sigma_{R.P. + L.T.}$

$$= 39.2(2.25 + .75) + 1.64[39.2(2.25 + .75)]^{1/2}$$

$$= 117.6 + 17.9 = 135.5 \text{ units}$$

(We have again assumed that the demand distribution can be adequately represented by the normal approximation of the Poisson distribution.)

(4) And from the above, the reserve stock = 17.9 units. The flow of computations which parallel the diagram in Fig. 8-10 are shown in the accompanying table.

TABLE OF P-MODEL SIMULATION
WITH RANDOM DEMAND AND RANDOM LEAD TIME SERIES

MONTH	S.O.H.	ISSUES	ORDERS	RECEIPT
0	106.0			
1	66.0	40.0 40		
*1.75	28.5	37.5}	107.0	
2	16	12.5}		
2.50	1.0-108.0	15.0}		107.0
3	93.0	15.0}		
*4	48.0	45.0 45	87.5	
4.50	30.5-118.0	17.5}		87.5
5	100.5	17.5}		
6	55.5	45.0		
*6.25	43.0	12.5}	92.5	
6.50	30.5-123.0	12.5}		92.5
7	98.0	25.0}		
8	58.0	40.0 40		
*8.50	38.0	20.0}	97.5	
9	18.0	20.0}		
9.25	6.75-104.25	11.25}		97.5
10	70.5	33.75}		
*10.75	48.0	22.5}	87.5	
11	40.5-128.0	7.5}		87.5
12	88.0	40.0 40		

We see plainly enough that the general situation has deteriorated but that there is nothing at all that we can do about this. It is nature that is extracting its payment. The increase in the variability, as a result of the lead time generator being hooked into the system, causes some periods in which a stock-out almost results, while in other periods the inventory level is higher than one would normally like. Judging from the results of our very limited simulation, a value for $L \cdot T_{\max}$ of three-quarters of a month is just about right. As previously stated, we could try other values if we thought we could do better. Certainly we would reach no final decision on the basis of a twelve-month simulation. Twenty years might be more like it. But the essence of

what we are talking about is conveyed by even a twelve months' sample. We will sum up for this simulation by looking at the moe which are:

B: Running average inventory in units = 69.1.

C: Cumulative number of orders = 5.3.

D: Running carrying cost of average dollar inventory = \$33.17.

E: Cumulative order cost = \$21.20.

F: Total cost = \$54.37.

In view of the low probability lead time series which was generated by chance, our rule box appears to be quite effective in coping with the existing situation. Of course, even worse situations are sure to arise—if we wait long enough.

68. SIMULATION OF INVENTORY QUEUING MODEL

We have been previously familiarized with the adaptation of a queuing model for low demand-high outage penalty items. Here again, our systems analysis and the resulting simulation can prove of great benefit. Consider the following simple problem: An airline repair center experiences an expected demand for jet engine replacements of two such engines per month. The demand distribution is Poisson with mean, λ of two units per month. When the repair center is out of stock the expected lead time $1/\mu$ to obtain an engine replacement is one-half month. The replenishment time distribution function which denotes the probability that the lead time will be longer than a time, t , is of the exponential type. From this we derive the pure number, ρ , which is the utilization factor of our queuing model, assuming that x units are carried in stock.

$$\rho = \frac{\lambda}{x\mu} = \frac{(2)\left(\frac{1}{2}\right)}{x} = \frac{1}{x}$$

Assume that one time in 20 years a shortage is to be allowed. The expected demand in a 20-year period would be:

$$20 \text{ years} \times 12 \text{ months/year} \times 2 \text{ units/month} = 480 \text{ units}$$

Thus we require protection against more than 1 outage in 480 demands. P_o is the probability of an outage per demand,* so $P_o = 1/480 = .002$. For the theoretical analysis we solve the queuing equation for various values of x ,

* See page 149 for the appropriate mathematical statement.

which represents the number of items to be maintained in stock, until our P_o criterion is realized: Thus,

$$x = 1 \quad P_o = (1)^{-1} = 1$$

$$x = 2 \quad P_o = [1 + (1 + 1)]^{-1} = \frac{1}{3}$$

$$x = 3 \quad P_o = \left[1 + 2 \cdot 2 \left(1 + 1 + \frac{1}{2}\right)\right]^{-1} = \frac{1}{11}$$

$$x = 4 \quad P_o = \left[1 + 3!3 \left(1 + 1 + \frac{1}{2} + \frac{1}{6}\right)\right]^{-1} = \frac{1}{49}$$

$$x = 5 \quad P_o = \left[1 + 4!3 \left(1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24}\right)\right]^{-1} = \frac{1}{261}$$

$$x = 6 \quad P_o = \left[1 + 5!4 \left(1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120}\right)\right]^{-1} = \frac{1}{1631}$$

Therefore, if stock on hand plus stock on order equals six units at all times, we expect to meet and pass our criterion.

Now let us hook our symbolic demand and lead time generators into our rule box which is:

- (1) S.O.H. + S.O.O. = 6 units.
- (2) Place an order to replace each issue as each issue occurs.

Since the expected demand per month is equal to two units and follows the Poisson distribution, we will obtain the expected demand per week and utilize the Poisson distribution of weekly requests. The demand generator will then sample the following distribution—which has mean and variance equal to .50. (We assume 4 weeks per month and that the demand always occurs in the middle of the week. If the assumptions are too inaccurate or restrictive they can be modified at the expense of additional computation.)

NUMBER OF ISSUES PER WEEK	PROBABILITY OF i ISSUES	CUMULATIVE PROBABILITY	MONTE CARLO NUMBERS
0	.607	.607	000-606
1	.303	.910	607-909
2	.076	.986	910-985
3	.012	.998	986-997
4	.002	1.000	998-999

Using our random numbers as before, we will obtain the following demand series covering a half-year period of time.

WEEK NUMBER	RANDOM NUMBER	DEMAND	L.T. (WEEKS)
1	470	0	..
2	881	1	1
3	121	0	..
4	684	1	0
5	388	0	..
6	833	1	4
7	845	1	0
8	472	0	..
9	647	1	0
10	728	1	1
11	129	0	..
12	553	0	..
13	403	0	..
14	157	0	..
15	902	1	1
16	618	1	3
17	870	1	1
18	982	2	2-0
19	099	0	..
20	588	0	..
21	116	0	..
22	610	1	0
23	118	0	..
24	022	0	..
25	636	1	1
26	726	1	4
		14	

Total demand equals 14, as shown. The last column gives the lead times which are to be associated with the order that is routinely placed when an engine is demanded and taken from stock. The lead time column was obtained by methods which should be "old hat" by now. Namely the lead time generator samples the replenishment time distribution which is described by $P(L.T. \geq t) = e^{-\mu t} = e^{-2t}$. Accordingly,

t (MONTHS)	e^{-2t}	$P(t)$	CUMULATIVE $P(L.T. \leq t)$	M.C. NUMBERS
0	1.000	.393	.393	000-392
1/4	.607	.239	.632	393-631
1/2	.368	.145	.777	632-776
3/4	.223	.088	.865	777-864
1	.135	.053	.918	865-917
5/4	.082	.032	.950	918-949
3/2	.050	.020	.970	950-969
7/4	.030	.012	.982	970-981
2	.018	.007	.989	982-988
9/4	.011	.004	.993	989-992
5/2	.007	.003	.996	993-995
11/4	.004	.002	.998	996-997
3	.002	.002	1.000	998-999

There were 14 issues and, correspondingly, 14 orders placed and, necessarily, 14 individual lead times. We continue generating random numbers and in this way we obtain the lead time column of the next table, which is also shown above.

ISSUE NUMBER	RANDOM NUMBER	LEAD TIME (WEEKS)
1	493	1
2	167	0
3	879	4
4	065	0
5	293	0
6	534	1
7	399	1
8	781	3
9	511	1
10	765	2
11	222	0
12	263	0
13	465	1
14	896	4

(We assume that 0 means that a delivery of the unit is made almost immediately.) Figure 8-11 illustrates the result of this simulation.

We see that by specifying such a very high protection level, the stock on

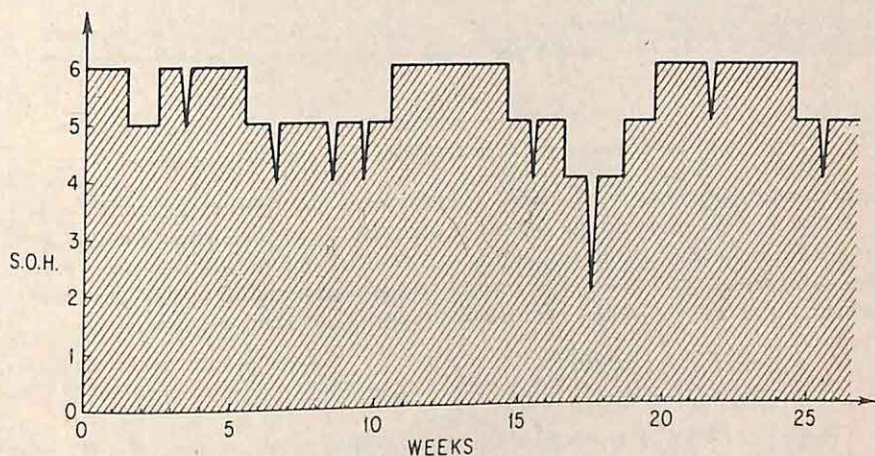


FIG. 8-11 Simulation of the inventory queuing model.

hand has not dropped below two units in the half-year period. Consulting our previous calculations for the determination of stock level for the queuing model, we find that if we had chosen $x = 3$ units, $P_0 = 1/11$. This means one outage every eleven demands. Two demands occur per month; therefore, eleven demands are expected in $5\frac{1}{2}$ months. This is roughly the period

which is covered by our simulation. It is apparent that if we had carried three instead of six units, we would have experienced the expected outage.

As a matter of interest, let us compare this queuing result with that of a Q -model designed to be analogous. The reorder point is the sum of the demand in the lead time period and the reserve stock. We will take L.T. equals one-half month. Then $\bar{z}_{L.T.} = 1$ unit and $\sigma_{z_{L.T.}} = \sqrt{1} = 1$. With this Poisson distribution, when $K = 4.0$, we obtain a protection level

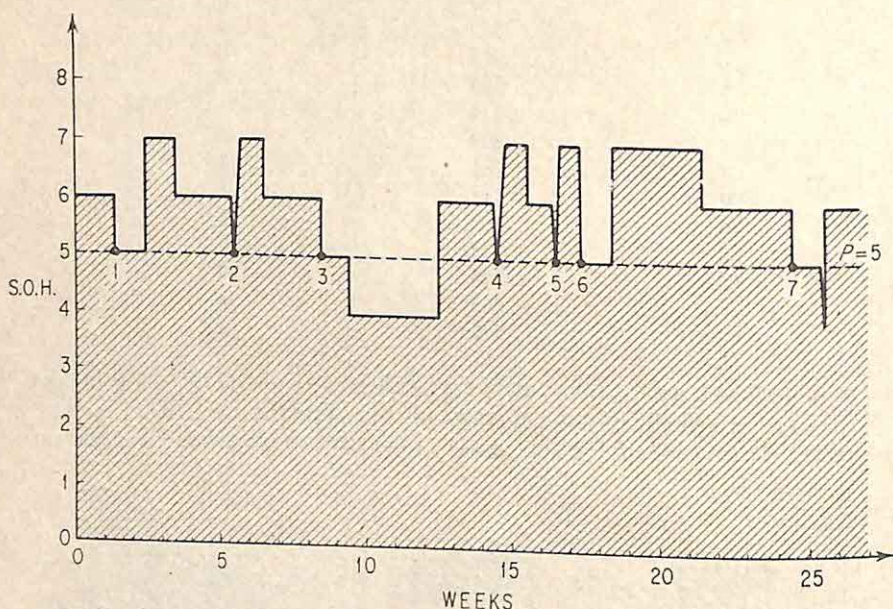


FIG. 8-12 Simulation of the Q -model analog to the queuing model.

approximately equivalent to .002. This gives us the reorder point: $P = 1 + 4(1) = 5$. We must design the order quantity, x , so that when one unit is demanded the reorder point is reached. A reorder quantity of two units will satisfy this requirement, one unit being the expected demand in the lead time interval. For $x_o = 2$, and 13 four-week months per year, we have:

$$2 = \sqrt{\frac{2(26)C_r}{cC_c}} \quad \text{and} \quad \frac{4}{52} = \frac{1}{13} = \frac{C_r}{cC_c}$$

Let $C_c = .10$ per year, then: $c/C_r = 130$, which gives us a rough idea of how much greater the cost per unit of the item must be than the order cost for this system to provide an optimal result. Let us now see how our Q -model queuing model. The expected number of orders per year is equal to $26/2 = 13$ per year or 6.5 semiannually. As each order is placed we will associate it with the same sequence of randomly generated lead times as was formerly used. Figure 8-12 is the diagram of the simulation. We see that

seven orders were placed, slightly exceeding our expectations. While the number of orders relative to the queuing system has been reduced, the average inventory investment appears to have increased. It is inadvisable to rely upon our eyesight in this matter so we compute the six-month moe and we find:

	QUEUING MODEL	Q-MODEL
B:	5.42 units	5.88 units
C:	14 orders per half-year	7 orders per half-year
D:	$ScC_c(5.42)$	$ScC_c(5.88)$
E:	$SC_r(14)$	$SC_r(7)$
F:	$(5.42)cC_c + 14C_r$	$(5.88)cC_c + 7C_r$

The *Q*-model has a slightly higher average inventory and a lower order frequency. The two models are equivalent when:

$$5.42cC_c + 14C_r = 5.88cC_c + 7C_r \quad (\text{half-year total costs})$$

Letting $C_c = .05$, since only a six-month period is covered, we find:

$$c/C_r = 304$$

Thus, for these data the queuing model is preferable when $c/C_r > 304$ and the *Q*-model is preferable when $c/C_r < 304$. This at least establishes a dividing line for us in terms of these two models. It also emphasizes the point that queuing models of this single-order type can only be used for high price items. All of which naturally leads to the question, "What about the *P*-model?" Let us frame the *P*-model as follows:

(1) The review period is one week.

(2) $L.T_{\max} = E(L.T.) = \frac{1}{2}$ month = 2 weeks.

(3) $M = \bar{z}_{R.P. + L.T.} + K\sigma_{R.P. + L.T.} = 2\left(\frac{1}{4} + \frac{1}{2}\right) + K\left[2\left(\frac{1}{4} + \frac{1}{2}\right)\right]^{1/2}$

$$= 1.5 + K\sqrt{1.5}$$

(4) When $K = 3.7$, $\alpha \approx .002$.

(5) Therefore, $M = 6.0$ units.

(6) And reserve stock = 4.5 units.

(7) There will be 52 review periods per year.

(8) But the expected number of orders per year, $n = 26$.

(9) Thus $x_o = z/n = 26/26 = 1$ unit.

Simulation of this *P*-model will produce exactly the same results as those we obtained from the queuing model. They are identical and Fig. 8-11 applies equally to both. Let us then find that value of c/C_r which would produce optimal results for both the *P* and queuing models. This is easily derived since,

$$x_o = 1 = \sqrt{\frac{2(26)C_r}{cC_c}}$$

and letting $C_c = .10$ we obtain $c/C_r = 520$. Comparing these c/C_r values we recall:

Q -model optimal: $c/C_r = 130$

Q , P and queuing model are equivalent—based on the simulation result: $c/C_r = 304$

P and queuing model optimal: $c/C_r = 520$

This is an example of some of the things we can learn from simulation. It must be remembered that because of the variable lead time factor, an analytical solution was avoided. Longer simulation runs should be required and then, on the basis of an adequate sample, relevant comparisons between models can be made. Many other factors can be introduced and varied such as discounts, expediting costs, and price changes. They are comparatively simple to include in the simulation although many times they are forbiddingly complex in mathematical form.

69. THE MANY ITEM SIMULATION

We have already introduced the notion that individual item optimization is almost guaranteed to destroy the collective, system-wide optimization. This important systems concept was explained in terms of the many item, the many location and the many department effects. We will now consider a simulation based on the many item effect. Because of the complexity of this system a flow diagram is particularly useful. Figure 8-13 supplies the map we require.

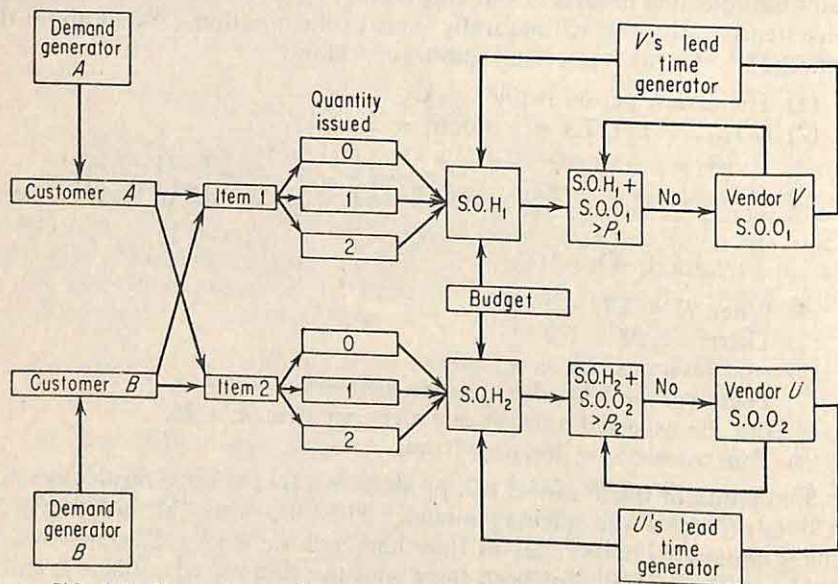


FIG. 8-13 Flow diagram of the many item simulation.
 S.O.O, stock on order for j th item
 S.O.H, stock on hand for j th item
 P , reorder point for j th item

We can read this flow diagram directly. It tells us how things work and what must still be specified. For the sake of simplicity the number of customers is limited to two. Similarly, the number of items is fixed at two and the demand quantities are limited to 0, 1, and 2 units. Only two vendors are used. Each has his own lead time characteristics. We will assume constant lead times:

Lead time of vendor V for item 1 = 2 weeks

Lead time of vendor U for item 2 = 4 weeks

Further, item 1 sells for \$140 per unit; item 2 sells for \$210 per unit. And, item 1 costs \$100 per unit; item 2 costs \$150 per unit. Let $C_c = .10$ per year.

The demand generator will produce an A and B sample for each day. So that it is possible to follow this example with relative ease, each customer is permitted to buy either quantities of item 1 or of item 2 on any one day—but not both. Then:

CUSTOMER	ITEM	QUANTITY	PROBABILITY	M.C. NUMBERS
A	none	0	.40	00-39
A	1	1	.20	40-59
A	1	2	.10	60-69
A	2	1	.20	70-89
A	2	2	.10	90-99
B	none	0	.50	00-49
B	1	1	.10	50-59
B	1	2	.05	60-64
B	2	1	.15	65-79
B	2	2	.20	80-99

From the above, for the purpose of obtaining the system's parameters and then designing appropriate rule boxes, we can obtain the distribution of total daily demand for each item:

A	B	$A + B$	$P(\text{ITEM 1})$	$P(\text{ITEM 2})$
0	0	0	$.70 \times .85 = .595$	$.70 \times .65 = .455$
1	0	1	$.20 \times .85 = .170$	$.20 \times .65 = .130$
0	1	1	$.70 \times .10 = .070$	$.70 \times .15 = .105$
1	1	2	$.20 \times .10 = .020$	$.20 \times .15 = .030$
2	0	2	$.10 \times .85 = .085$	$.10 \times .65 = .065$
0	2	2	$.70 \times .05 = .035$	$.70 \times .20 = .140$
2	1	3	$.10 \times .10 = .010$	$.10 \times .15 = .015$
1	2	3	$.20 \times .05 = .010$	$.20 \times .20 = .040$
2	2	4	$.10 \times .05 = .005$	$.10 \times .20 = .020$
			1.000	1.000

Which reduces to:

DEMAND PER DAY ITEM 1	PROBABILITY ITEM 1	DEMAND PER DAY ITEM 2	PROBABILITY ITEM 2
0	.595	0	.455
1	.240	1	.235
2	.140	2	.235
3	.020	3	.055
4	.005	4	.020

the expected demand per day for item 1 is .60 units. The expected demand per day for item 2 is .95 units. We have been told that item 1 costs \$100 per unit and item 2 costs \$150 per unit. Now we must raise two conditions of great importance. First, the company wishes to limit its inventory investment to no more than \$2500 invested in the combined average inventory of the two items. (Note the budget limitations in Fig. 8-13.) Reorder points designed to give about 95% protection are calculated; i.e., $P_1 = 6 + 2\sqrt{6} = 11$ units and $P_2 = 19 + 2\sqrt{19} = 28$ units. For the first item the reserve stock is about five units, which is a yearly inventory investment of \$500. For the second item the reserve stock is nine units at \$150 per unit or \$1350. The combined investment is then \$1850, which leaves only \$650 to be spent on the combined average order quantities; i.e.,

$$\frac{c_1 x_1}{2} + \frac{c_2 x_2}{2}$$

Ordinarily we would determine x_0 for each item and proceed on our merry way. Thus, if the cost per order for either item was \$5.00 and a business year consists of 240 days:

$$x_0(1) = \left(\frac{2 \times .60 \times 240 \times 5}{100 \times .10} \right)^{1/2} = 12 \text{ units}$$

and

$$x_0(2) = \left(\frac{2 \times .95 \times 240 \times 5}{150 \times .10} \right)^{1/2} = 12.3 \text{ units}$$

Hence,

$$\frac{[100(12) + 150(12.3)]}{2} = \$1523$$

This is over \$800 in excess of the budgeted amount.

A number of things can be done. We can reduce reserve stocks in either item or both. We can order smaller quantities of either item or both. But what should we do? The systems concept can be applied by systematically experimenting with various approaches. Before we begin we will complicate our situation with the second condition which was not explained but mentioned above. Namely, an outage experienced by customer *A* shifts all his business away for a two-week period. If experienced by customer *B* it results in his buying everything from a different source for three weeks.

TABLE FOR MANY ITEM SIMULATION I

RANDOM NUMBERS		DAY	DEMAND				ITEM	
A	B		A		B		1	2
			1	2	1	2		
47	57	1	1	0	1	0	2	0
08	90	2	0	0	0	2	0	2
81	26	3	0	1	0	0	0	1
12	18	4	0	0	0	0	0	0
16	87	5	0	0	0	2	0	2
84	09	6	0	1	0	0	0	1
38	82	7	0	0	0	2	0	2
88	09	8	0	1	0	0	0	1
33	95	9	0	0	0	2	0	2
84	88	10	0	1	0	2	0	3
54	11	11	1	0	0	0	1	0
72	66	12	0	1	0	1	0	2
64	10	13	2	0	0	0	2	0
77	11	14	0	1	0	0	0	1
28	80	15	0	0	0	2	0	2
12	22	16	0	0	0	0	0	0
95	63	17	0	2	2	0	2	2
53	67	18	1	0	0	1	1	1
40	26	19	1	0	0	0	1	0
31	14	20	0	0	0	0	0	0
08	19	21	0	0	0	0	0	0
69	34	22	2	0	0	0	2	0
28	50	23	0	0	1	0	1	0
72	34	24	0	1	0	0	0	1
37	69	25	0	0	0	1	0	1
87	67	26	0	1	0	1	0	2
87	92	27	0	1	0	2	0	3
90	77	28	0	2	0	1	0	3
16	68	29	0	0	0	1	0	1
29	55	30	0	0	1	0	1	0
72	70	31	0	1	0	1	0	2
18	27	32	0	0	0	0	0	0
43	85	33	0	0	0	2	1	2
38	55	34	1	0	1	0	1	0
53	97	35	0	0	0	0	1	2
95	44	36	1	0	0	2	1	2
14	27	37	0	2	0	0	0	2
46	88	38	0	0	0	0	0	0
79	09	39	0	0	0	2	1	2
43	40	40	1	0	0	0	0	1
			1	0	0	0	1	0
			12 17		6 27		18 44	
			29		33		62	

The first set of rule boxes to be tried will be: $P_1 = 11$, $x_1 = 5$, $P_2 = 28$, $x_2 = 5$. This meets the budget restriction. Let us see how it affects the company's position, bearing in mind the limited sample size that we will use.

First, we let our random generator produce two months supply of input data. The previous table with the Monte Carlo numbers is used to derive the accompanying simulation table.

Figure 8-14 presents the simulation based on the inventory strategy, $P_1 = 11$, $P_2 = 28$, $x_1 = 5$, $x_2 = 5$. We find that item 2 requires time to

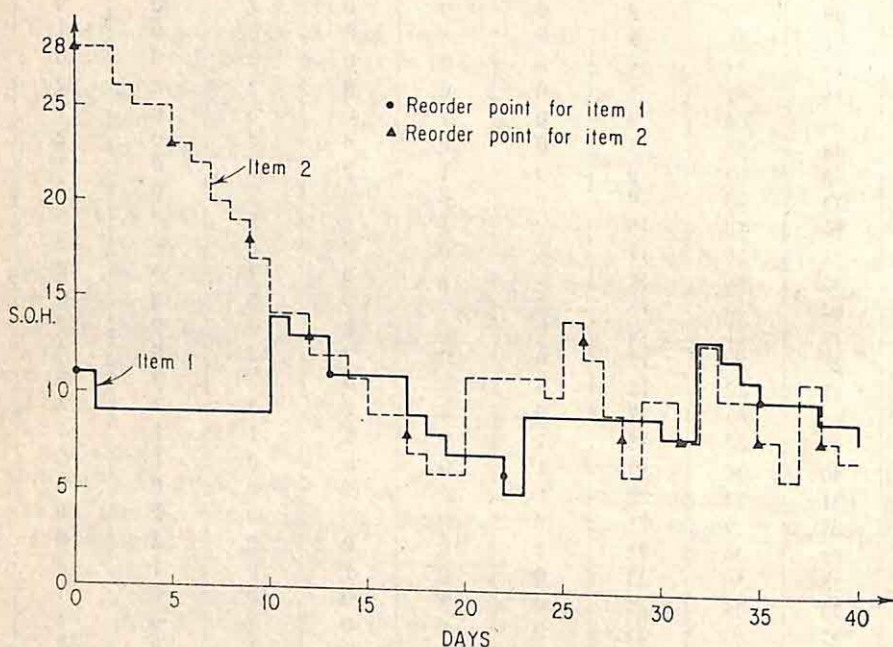


FIG. 8-14 Many item simulation I.

settle down. It continually decreases until the orders which have been placed begin to take effect. Such initial phenomena as oscillations, rapid growth, and rapid decay are not unusual in simulations. A sufficient period should always be allowed for stability to set in. Even so, in real life a system has to be introduced at a point in time and the system actually experiences many forms of transients. It is useful then to be able to simulate these initial conditions for at least three reasons. First, so that deteriorations which occur immediately after a new system has been introduced should not be misinterpreted. If they are anticipated no emergency sessions will be called to cancel the program for the new system which is being installed. Second, so that we are able to change the system's design if undesirable initial conditions are observed. After all, the initial conditions are a means to an end and there is little point in allowing the means to destroy the possibility of ever reaching the end. Third, mathematical analysis of system transients is

forbidding. Most of our strong analytical methods are based on the system achieving a steady state. As in this case, it can take a while.

We will calculate relevant moe beginning with the twentieth day so as to eliminate the transient responses. But the transient response patterns are available if the need exists to use them. We find for plan 1, which is our first strategy:

moe(MONTHS)	ITEM 1	ITEM 2	TOTAL
B	9.1	9.8	19
C	2	5	7
D	\$7.58	\$12.25	\$19.38
E	\$10.00	\$25.00	\$35.00
F	\$17.58	\$37.25	\$54.83
Sales volume	9	22	31
Sales cost	\$900.00	\$3300.00	\$4200.00
Total cost	\$917.58	\$3337.25	\$4254.83
Dollar volume	\$1260.00	\$4620.00	\$5880.00
Profit	\$342.42	\$1282.75	\$1625.17
Profit/unit	\$38.05	\$58.31	\$52.42
Stock-out cost	none	none	none
F/unit	\$1.95	\$1.69	\$1.77
F/dollar-unit	.020	.011	.013

The record looks better for item 2 in spite of the fact that it has a higher total cost. We must remember that we are comparing the effect on two different items of a single strategy. Therefore, we have divided the F -measure (which is the total cost), by the number of units demanded. We have also divided the F -measure by the dollar value of the units demanded. Both measures indicate that item 2 is performing more efficiently. At the same time, a higher profit per unit is derived from item 2.

Obviously, we can work through many different feasible* simulation programs. A few such feasible programs would be:

	K_1	K_2	x_1	x_2	
Plan 1	2	2	5	5	(FIG. 8-14)
Plan 2	1	2.4	5	5	
Plan 3	4	1.93	4	12	(FIG. 8-15)
Plan 4	0	2.60	4	15	(FIG. 8-16)
Plan 5	2	2	7	4	

Consider the second plan suggested above. Its effect is felt in the reorder point. The revised reorder points would be:

$$P_1 = 6 + \sqrt{6} \cong 9 \text{ units}$$

$$P_2 = 19 + 2.4\sqrt{19} \cong 30 \text{ units}$$

* By feasible we mean that the budget restriction:

$$(100)(\sqrt{6})K_1 + (150)(\sqrt{19})K_2 + 100x_1/2 + 150x_2/2 = 250K_1 + 675K_2 + 50x_1 + 75x_2 = \$2500$$

is satisfied.

Since no stock-outs result, at least with the sample obtained from this particular run of our demand generator, we can evaluate the results by lowering the item 1 curve by two units everywhere and by raising the item 2 curve by two units everywhere. The revised moe for plan 2 will be:

moe(MONTH)	ITEM 1	ITEM 2	TOTAL
B	7.1	11.8	19
C	2	5	7
D	\$5.92	\$14.75	\$20.67
E	\$10.00	\$25.00	\$35.00
F	\$15.92	\$39.75	\$55.67
Sales volume	9	22	31
Sales cost	\$900.00	\$3300.00	\$4200.00
Total cost	\$915.92	\$3339.75	\$4255.67
Dollar volume	\$1260.00	\$4620.00	\$5880.00
Profit	\$344.08	\$1280.25	\$1624.33
Profit/unit	\$38.23	\$58.19	\$52.40
Stock-out cost	none	none	none
F/unit	\$1.77	\$1.81	\$1.80
F/dollar-unit	.018	.012	.013

This plan does not seem to offer any advantage since it has a slightly higher total cost and a commensurately lower profit. Of course, it must be remembered that over the long run we should observe more item 1 stock-outs and

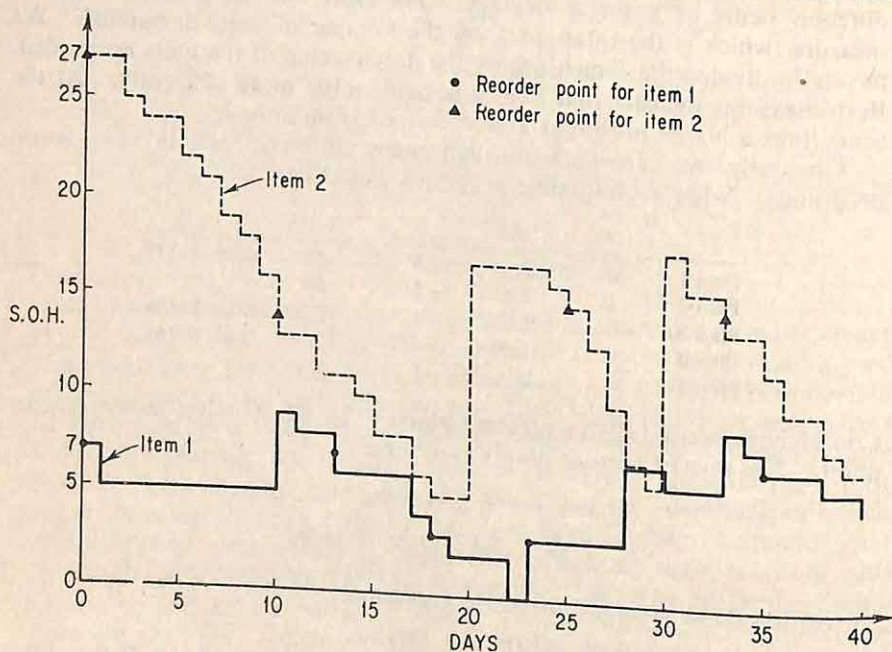


FIG. 8-15 Many item simulation II.

fewer item 2 stock-outs with the second plan than with the first plan. Since each kind of stock-out does not have the same effect—it is possible that the second plan would, in fact, be preferred. With these thoughts in mind, let us simulate two more plans.

Plan 3: $P_1 = 7$, $P_2 = 27$, $x_1 = 4$, $x_2 = 12$ (Fig. 8-15)

Plan 4: $P_1 = 6$, $P_2 = 31$, $x_1 = 4$, $x_2 = 15$ (Fig. 8-16)

Both plans satisfy the budget restriction of \$2500. But we have introduced another complication; plan 4 takes advantage of a discount which is available for item 2. Namely, with orders of 15 units or more the price per unit is reduced to \$120. Figures 8-15 and 8-16 graphically portray the

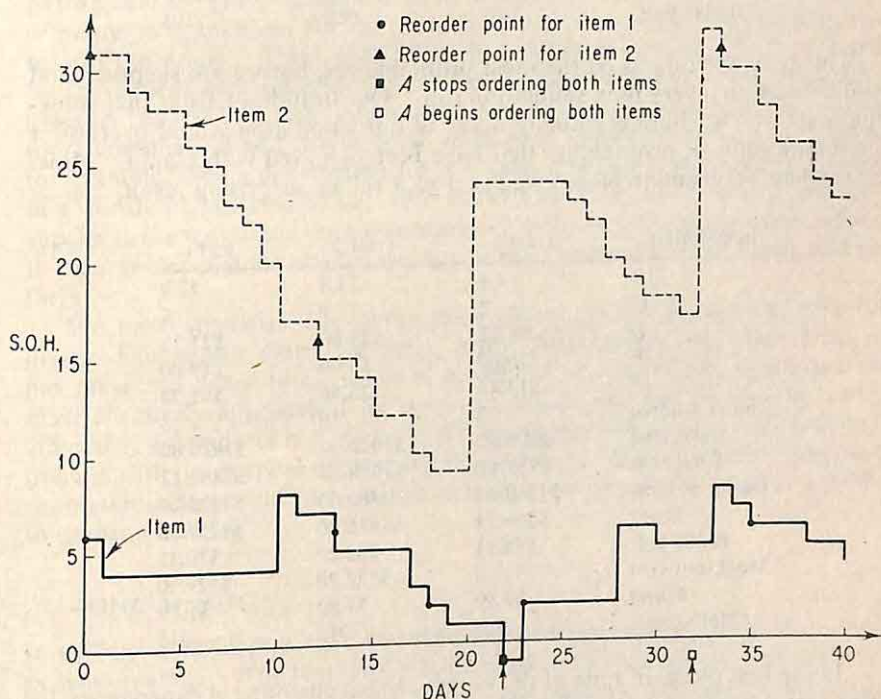


FIG. 8-16 Many item simulation III.

respective situations. In the former, the stock level of item 1 is zero during the twenty-third day, but no stock-out results and a reasonable recovery is made after that. Plan 4 (Fig. 8-16) produces a stock-outage. Reference to our tables shows that it is customer A who experiences the shortage and becomes incensed. This occurs between the twenty-third and the thirty-third day during which time he shifts all his business to a competing firm. Let us see what happens to the plan 3 move, again using only the data of the second month.

moe(MONTH)	ITEM 1	ITEM 2	TOTAL
<i>B</i>	5.1	13.0	18.1
<i>C</i>	2	2	4
<i>D</i>	\$4.25	\$16.25	\$20.50
<i>E</i>	\$10.00	\$10.00	\$20.00
<i>F</i>	\$14.25	\$26.25	\$40.50
Sales volume	9	22	31
Sales cost	\$900.00	\$3300.00	\$4200.00
Total cost	\$914.25	\$3326.25	\$4240.50
Dollar volume	\$1260.00	\$4620.00	\$5880.00
Profit	\$345.75	\$1293.75	\$1639.50
Profit/unit	\$38.42	\$58.81	\$52.89
Stock-out cost	none	none	none
<i>F</i> /unit	\$1.58	\$1.19	\$1.31
<i>F</i> /dollar-unit	.016	.008	.010

This plan appears to be the most profitable yet, but we are skeptical and would insist on a very long simulation run. Our intuition tells us that stock-out costs which would eventually occur in our simulation would overpower the minor gains in profitability that have been achieved with plan 3. However, when we examine plan 4 we discover a rather surprising result.

moe(MONTH)	ITEM 1	ITEM 2	TOTAL
<i>B</i>	4.1	23.8	27.9
<i>C</i>	2	1	3
<i>D</i>	\$3.42	\$23.80	\$27.22
<i>E</i>	\$10.00	\$5.00	\$15.00
<i>F</i>	\$13.42	\$28.80	\$42.22
Sales volume	9	16	25
Sales cost	\$900.00	\$1920.00	\$2820.00
Total cost	\$913.42	\$1948.80	\$2862.22
Dollar volume	\$1260.00	\$3360.00	\$4620.00
Profit	\$346.58	\$1411.20	\$1757.78
Profit/unit	\$38.51	\$88.20	\$70.31
Stock-out cost	0	\$529.20	\$529.20
<i>F</i> /unit	\$1.49	\$1.80	\$1.69
<i>F</i> /dollar-unit	.015	.015	.015

In the first place, in spite of the outage, which customer *A* experiences, the profit level has risen to the highest point of all of the four plans. This occurs even though an outage occurs. The cost of this outage in foregone profit is \$529.20. It is apparent that the discount for purchasing item 2 in quantities of 15 or more units is responsible. The effect of the discount is noticeable in the increased per unit profit of item 2. Meanwhile, the per unit profit of item 1 remains essentially unchanged.

The stock-out situation occurs when customer *A* places an order for two units of item 1 at the beginning of the twenty-third week. The first unit is immediately delivered. The second unit is back-ordered and shipped the next day to customer *A*. But *A*, according to our rules, has been lost as a

customer of the company for a period of 10 days. Now, an interesting fact is that if *A* had continued to buy from our company, taking no notice of the stock-outage, he would have bought no units of item 1, but he would have bought six units of item 2. Thus, although the outage occurred as the result of a shortage of item 1, item 1 suffers no penalty whatever. Item 2 is the loser. The usual inventory approach which ignores the total system of dependencies would penalize only item 1. No thought would be given to the effect on item 2.

It would seem that this example drives home our point about the interdependencies of systems and the need for systems analysis more effectively than a thousand words. Other interesting facets of the simulation are available from the movie tables. They reinforce the fact that complex situations can be very difficult to intuit. Bounds on rationality lead one naturally to fragment the problem and find good solutions for the parts. But the sum of many good solutions for the parts does not in any way foretell the goodness of the whole. Think then about how much complexity can be treated by the patient and detailed approach of systems analysis. We must also consider the fact that this system which we have used might have been much more complicated and it still could have been simulated. We could have, for example, used a distribution for lost sales resulting from outages instead of a constant, expected value. We could have considered various vendors supplying the same item, the advantages of expediting orders, ways and means to recover lost customers more quickly, many customers, many items, and so forth.

We have simulated the many item effect. Other simulations can be designed to handle many locations, many departments, etc. Our chapter has presented just a brief review of a few different problems which can be most effectively approached by systems analysis and the simulation route. Obviously a computer is a big help in running through a variety of company policies with the purpose of selecting the best one. But even hand computed simulations can provide worthwhile benefits if properly conceived as a guide to decision-making.

70. CONCLUSION

It will be remembered that we dichotomized our approach to inventory systems analysis. The first path was directed to the investigation of decision rules. We have now followed along that path a reasonable distance. Reasonable, that is, in terms of our purposes and the inherent limitations of a textbook. We have tried to demonstrate that while the format of our systems analyses remains the same, the breadth of its possibilities is enormous. Ingenuity must be called upon to devise the program of experimentation. Very exact and detailed simulations of poorly conceived rule boxes may enable us to choose the best from a totally undesirable set of possibilities. Intuition and judgement can help to eliminate those experiments which would be a waste of time. Analytical aids can bolster such judgement and frequently can spark the native creativity of the decision-maker.

There should not be confusion here. Simulation and analytical methods of solution work hand in hand. When the latter is available the former can be used to test the conclusions under circumstances which are closer to real life. No analytical solution is free from premises and assumptions. The simulation can be used to observe the effects obtained when some of these assumptions are relaxed. When no analytical solution is possible the program of experimental simulation can be used. While such procedures do not even hold the promise of producing an optimal or best possible solution they do lead us in the right direction. Meaningful results can be obtained, frequently, by holding all factors but one constant and then varying that one until a best point is reached. Using that best point, another factor is then varied. If interaction exists it will sometimes be apparent. In any case, the rules for intelligent experimentation apply here as well as in the physical laboratory.

Simulation is only one aspect of the systems analysis approach. Analytical methods exist as well and we shall be considering some of these as we continue. To a great extent the use of simulation is a nonrecurring cost factor in the inventory study. It is also a sizeable component of such costs. The justification for such expenses should be evident by now. But there are additional benefits which result directly from the other parts of our inventory study.

PROBLEMS

1. In the text, a dynamic programming matrix was developed and set forth. Two items were involved, P_1 and P_2 . The objective was to minimize the total cost, TC, of these two items, subject to a budgetary restriction. Extend this example to three items as follows:

	c	C_c	C_r	z
P_1	\$2.00	.10	\$4	10
P_2	5.00	.10	4	16
P_3	20.00	.10	4	4

(a) If the budget allows an average dollar inventory of \$100, what is the optimal allocation of the budget among the items?

(b) If the budget is fixed at \$50, what is the optimal allocation among the items?

2. Construct a master block flow diagram for the materials management function. Describe as many components as you can. Then, on the basis of the master, segment and expand components—such as purchasing and inventory management—by detailing their flow and decision systems.

3. Design a demand generator to simulate a horizontal distribution and a triangular distribution whose upper and lower bounds are 100 and 40 units per month. Using the same random number sequence (which appears

below) for both distributions, generate a sample of 25 demand values for each type of input. (Read left to right.)

(06	58	04	62	85	07	88	48	39	28	77	41	05
48	80	55	94	80	88	09	21	14	41	96	01)	

Comment on the demand series which are generated by observing the differences between the theoretical distributions and the sample results.

4. Feed the demands generated above into a properly designed Q -model and P -model. Comment on the comparative results obtained by simulation.

The following data apply:

c = \$14 per item. α = .10.

C_r = \$5 per order. L.T. = 1 month.

C_c = .02 per month.

5. How will the queuing model fare?

6. Draw a flow diagram for the simulation of an Ss model (see text, page 115-116).

7. Work out the results that would be obtained if plan 5 (as presented in Section 69 were to be used.

the data processing study: design of operating procedures

Our concern in the previous chapter was with the design of satisfactory inventory rules. We now turn to the problem of designing a satisfactory operating system. We shall show that these two areas cannot be treated as being independent of each other. A variety of systemic costs tie these study phases together. Since both the decision rules and the data system represent a single functional entity, the systems approach is the only approach which can be trusted when a nontheoretical inventory study is being undertaken.

71. THE COST OF AN ORDER

In our inventory models we rather blithely talked about the cost of an order. We called this C_r and it appeared in the Q -model and in the P -model. We also showed that rational inventory policies were available even though the order cost was unknown. The cost per order was assumed to be very small, or at least relatively so, with respect to the queuing model. Everything that was said was reasonable as far as it went—but some things were left unsaid. We now propose to remedy that deficiency.

What is the order cost? It is composed of:

- (1) the cost of handling issue transactions,
- (2) the cost of handling receipt transactions,
- (3) the cost of making and sending the purchase order,
- (4) the cost of expediting,
- (5) the cost of updating,
- (6) the administrative and overhead costs,
- (7) other costs.

The question then arises, are these component costs independent of the quantity ordered? After all, we use C_r to determine the optimal order quantity, x_0 . This is true of the Q -model. In the P -model we use C_r to obtain the optimal review period. Is C_r , in this case independent of the size of the review period? The answer is that C_r is not independent of x_0 in the first case and it is not independent of the optimal review period in the second case.

Various kinds of dependencies exist, but the one that we will concern ourselves with is the issue transaction cost. Rather than indulge in circumlocution let us set down a general form for the cost of an order.

$$C_r = V_r + F_r = \frac{x_0 G_1}{v} + F_r$$

where V_r = the variable cost per order,

F_r = the cost of handling a single issue transaction,

G_1 = the fixed costs per order,

v = the average number of units per issue transaction.

We can determine G_1 by using work sampling methods. Most of the above listed costs can be determined in this way. Each item would be characterized by its own v . For some products the average number of units withdrawn at a time might be a dozen; for others, one unit might be characteristic.

First, let us consider the Q -model. It is a perpetual inventory system. Therefore, each transaction, including issue transactions, are handled one by one as they arise. The total quantity issued between successive reorder points is x_0 . Thus, if the average number of units per issue transaction is equal to one, then x_0 different issue transactions must be processed. If, on the other hand, $v = x_0$, then only one issue transaction must be processed. Obviously, the cost per order varies with the size of x_0/v . But, this does not affect the determination of the optimal order quantity. The cost component, which increases linearly with x , produces a constant, the derivative of which is zero. That is,

$$\begin{aligned} \frac{dT.C.}{dx} &= \frac{d}{dx} \left[\frac{cC_c x}{2} + \left(\frac{xG_1}{v} + F_r \right) \frac{z}{x} \right] = \frac{d}{dx} \left(\frac{cC_c x}{2} + \frac{zG_1}{v} + F_r \frac{z}{x} \right) \\ &= \left(\frac{cC_c}{2} - \frac{F_r z}{x^2} \right) = \left(\frac{cC_c}{2} - \frac{C_r z}{x^2} \right) = 0 \end{aligned}$$

So it is F_r which we have been calling C_r .

With the P -model, issue transactions can accumulate during the review period. At the review date they are gathered together and added to obtain the total quantity issued. The work load is obviously smaller since the stock card must be matched only once with the transaction slips and only a single entry has to be made. In spite of the lower cost per issue transaction of the P -model as compared to the Q -model, there is nevertheless a component of the order cost which varies with the number of transactions and the number of transactions which occur in a period will be directly related to the length

of the period. Therefore, our comments with respect to issue transactions and the Q -model can also be applied to the P -model.

Two conclusions can be drawn. First, the cost per order which we normally apply in deriving optimal order quantities and order periods is only part of the total order cost. It is the fixed part of the cost per order. Looked at another way, it is the variable component of the total order cost. Since this last point may cause a little confusion, let us restate the situation. Costs which are independent of the size of the order quantity and the number of orders per year; e.g., zG_1/v , are variable cost components of the cost per order and fixed components of the total costs. It is usual to name the cost component fixed or variable in terms of the role it plays in the total cost. Accordingly, it is clear that part of the order cost never appears in the ordinary inventory formulation. It isn't needed. But it certainly is a factor of importance since it appears in the total cost equation. Its importance lies in the comparison between systems.

Second, the P -model is less expensive to operate. The cost advantage of the P -model is hidden in the fixed costs which do not ordinarily appear in the inventory formulation. In fact, at first glance the P -model looks more expensive to operate since it is likely to have a slightly higher C_r (variable cost) than the Q -model. Since the kind of costs which do not appear are those which are invariant to order frequency or size, they must measure fixed responsibilities of the ordering system. The remainder of the total order cost must measure the variable responsibilities of the ordering system. Together, they measure the overall demands which the ordering system must satisfy.

72. PERPETUAL VS. PERIODIC DATA-PROCESSING

Look again at zG_1/v . z and v are characteristics of each item. For a more general representation we can add a subscript j . We then have: z_jG_1/v_j , where j stands for the j th item. And:

$$\sum_j \frac{z_j}{v_j} = \text{total number of issue transactions per year}$$

Now let G_1 equal the cost of handling an issue transaction with the Q -model and let G_2 equal the cost of handling an issue transaction with the P -model. The latter cost, G_2 , will be very much smaller than G_1 . It consists only of adding an issue quantity to a running total or entering the issue quantity on the keyboard of a machine. The cost of adjusting the stock card, of analyzing the results, etc., are variable costs depending upon the number of review periods in a year. These costs are then carried over to the F_r (or C_r) factor for the P -model. On the other hand, with the Q -model the F_r factor is smaller since the issue transaction costs per year are independent of the interval between orders. What effect does this have? Replacing F_r with our familiar C_r and referring to the economic lot-size formula,

$$x_o = \left(\frac{2zC_r}{cC_c} \right)^{1/2}$$

we observe that x_o will be larger for the P -model than for the Q -model, everything else being equal. Let us demonstrate this with a simple example. For the Q -model, assume that $z = 1200$ units per year, $v = 12$ units per issue transaction, $G_1 = \$0.10$ per transaction. (This includes finding the appropriate stock card, decreasing the stock on hand, and determining whether the reorder point has been reached. If the reorder point has been reached, it does not include the cost of the purchase order and other cost components that are a function of the number of orders placed.) Let \$2.00 equal the total variable order costs. Further, $c = \$4.00$ and $C_c = .12$. We then have:

$$x_o = \sqrt{\frac{2(1200)(2)}{4(.12)}} = 100 \text{ units}$$

and

$$\text{Total order cost} = \left[\frac{1200}{12} (.10) + \frac{2.00(1200)}{100} \right] = \$34$$

For the P -model everything remains the same except that $G_2 = \$0.01$ per transaction. (This includes only the cost of storing the transaction and adding it to other transactions.) However, we must increase the variable cost component to account for the procedure of locating the stock card, decreasing the stock on hand, etc. We will use the G_1 figure given above as a good approximation of this amount. We then have:

$$x_o = \sqrt{\frac{2(1200)(2.10)}{4(.12)}} = 102.5 \text{ units}$$

and

$$\text{Total order cost} = \left[\frac{1200}{12} (.01) + \frac{2.10(1200)}{102.5} \right] \approx \$25.60$$

There has always been a great deal of confusion about the meaning of the order cost and this discussion is intended to go to the heart of the matter to prevent such confusion from arising here. It is also intended to present the point of view that the major practical feature which differentiates one inventory model from another is the data-processing routine which is required by the model.

What happens, we might ask, if by mistake some fixed costs are included in the variable cost figure? Previously we have shown that errors in these estimates are considerably diluted and produce only relatively small changes in the total variable cost. This is true until we take a systemic point of view which requires consideration of both fixed and variable costs. Then, as we see above, the cost differences from a planning point of view can be very misleading if we do not take fixed costs into account. Using our example, we observe that the variable component of total order cost is \$24.00 for the Q -model and approximately \$24.60 for the P -model, a small increase which

causes a somewhat longer period to elapse between orders. That is,

$$t_Q = \frac{12}{n_Q} = \frac{12}{z} x_o^Q = 12 \sqrt{\frac{2C_r}{cC_c z}}$$

and

$$t_P = \frac{12}{n_P} = \frac{12}{z} x_o^P = 12 \sqrt{\frac{2(C_r + \Delta)}{cC_c z}}$$

whence

$$t_P > t_Q$$

where t_Q = the expected interval in months between orders using the Q -model,

t_P = the interval in months between orders using the P -model,

n_Q = the expected number of orders per year using the Q -model,

n_P = the number of orders per year using the P -model,

Δ = the additional variable cost resulting from the use of the P -model.

How can we interpret this result? We can interpret the longer interval between orders of the P -model as a net gain for the data-processing department. The gain is obtained as a result of an increased requirement for capital invested in the inventory. We know that the additional ordering cost, Δ , which appears in the P -model equation given above is something of a fiction since it actually represents a transfer of cost from the fixed to the variable category. It is implicit in the design of fixed order cycle systems that as many costs as possible will be converted to the variable type of cost component where they can be controlled. Control increases as we eliminate random demands on the processing system and instead regiment these demands to a previously specified schedule. The ordering costs which we can control may then be balanced against those opposing cost systems which are also under our control. *There is no point in balancing costs over which we have no control.* The cost of increased control can be traced to additional inventory investments of two kinds: (1) the increase which occurs in x_o when a fixed order cycle is used, (2) an increase in the reserve stock requirement. We will handle these separately. Our initial assumption, which we will later relax, is that the reserve stock requirements of the P - and the Q -model are sufficiently close so that they may be ignored.

Obviously, the company would not want to accept these additional costs unless an overall gain could be shown. The overall gain which results from increased data-processing control can be sizable, but it is by no means guaranteed. On the contrary, a substantial negative gain, or loss, can be the result. For our example, the total variable costs—including inventory carrying charges and order costs—increase from \$48.00 to about \$49.20 when we switch from the Q -model to the P -model. But this increase of \$1.20 is more than offset by the decrease in the fixed cost component of the total order cost, which changes from \$10.00 to \$1.00. The total benefit, ignoring additional reserve stock costs that results from increased control over the data-processing function, is \$7.80. That is,

Saving in fixed order costs	\$9.00
Increase in variable order costs	.60
Increase in inventory carrying costs	.60
Increase in total variable costs	<u>1.20</u>
Net savings (before reserve stock costs)	\$7.80

The percent savings to the company (before reserve stock costs) is about 13%, namely, $(\$58.00 - \$50.20)/\$58.00$. The percent savings of the data-processing department are naturally more substantial, approximately 25%, namely, $(\$34.00 - \$25.60)/\$34.00$. If this one item can be considered close to average for the total line of items that are carried by the company, then substantial savings might be available. But they might not be available, depending upon the reserve stock situation which we will now investigate.

The reserve stock effect is usually of considerable importance. It is non-existent only under very special circumstances. The *P*-model is obviously superior from the point of view of ordering and processing costs, and the savings to be derived from scheduling work are hardly ones to be missed by an alert executive. Then why should we ever use the *Q*-model? The answer is that, in general, the *P*-model requires a larger reserve stock than the *Q*-model does for the same level of service efficiency. Thus, the net saving of \$7.80 per item per year must be further reduced to account for the additional carrying costs that are attributable to reserve stock. The added reserve stock cost is:

$$cC_c K(\sigma_{R.P. + L.T.} - \sigma_{L.T.})$$

Therefore, depending upon the price of the item, the carrying cost rate, the service level required, the character of the demand distribution, the length of the lead time period (L.T.), and the length of the review period (R.P.), a net gain or loss can result. For our example, let us see under what conditions the saving of \$7.80 might be overshadowed by the additional reserve stock cost. We have:

$$4(.12)K(\sigma_{R.P. + L.T.} - \sigma_{L.T.}) = \$7.80$$

or

$$K(\sigma_{R.P. + L.T.} - \sigma_{L.T.}) = 16.25$$

Assume that $\sigma_{\text{year}} = 35$ units; whence,

$$\sigma_{\text{month}} = \frac{\sigma_{\text{year}}}{\sqrt{12}} = \frac{35}{\sqrt{12}} = 10 \text{ units}$$

Let L.T. = 1 month, R.P. = 1 month; therefore R.P. + L.T. = 2 months; then

$$\sigma_{R.P. + L.T.}^2 = 2\sigma_{\text{month}}^2; \quad \sigma_{R.P. + L.T.} = \sqrt{2} \sigma_{\text{month}}; \quad \sigma_{L.T.} = \sigma_{\text{month}}$$

and

$$(\sigma_{R.P. + L.T.} - \sigma_{L.T.}) = \sigma_{\text{month}}(\sqrt{2} - 1) = 10(.41) = 4.1 \text{ units}$$

From this we find that:

$$K = \frac{16.25}{4.1} \approx 4$$

The *K* value is quite high indicating that this is an item which requires unusual protection. Greater than average levels of protection is one of the conditions which leads the executive to favor the use of a perpetual inventory system; i.e., the *Q*-model family.

Now, let $L.T. = \frac{1}{2}$ month, and hold everything else fixed, as before. We obtain:

$$\sigma_{R.P. + L.T.} = \sqrt{1.5} \sigma_{\text{month}} = 12 \text{ units}; \quad \sigma_{L.T.} = \sqrt{.5} \sigma_{\text{month}} = 4 \text{ units}$$

$$\text{Then,} \quad (\sigma_{R.P. + L.T.} - \sigma_{L.T.}) = 5 \text{ units}$$

In this case,

$$K = \frac{16.25}{5} = 3.25$$

We see that the required protection has decreased and still the savings derived from more efficient clerical procedures are balanced against the reserve stock costs, so that we are indifferent as to whether we choose the Q - or the P -models. The factor which accounts for this is the greater size of the review period as compared to the lead time period. For the general case we derive:

$$\sigma_{R.P.}^2 = R.P. \sigma_{\text{month}}^2; \quad \sigma_{L.T.}^2 = L.T. \sigma_{\text{month}}^2$$

where $L.T.$ and $R.P.$ are given in months. Then,

$$\sigma_{R.P. + L.T.}^2 = (R.P. + L.T.) \sigma_{\text{month}}^2$$

$$\text{and} \quad (\sigma_{R.P. + L.T.} - \sigma_{L.T.}) = (\sqrt{R.P. + L.T.} - \sqrt{L.T.}) \sigma_{\text{month}}$$

Going back to our original balancing equation, we can now write:

$$K(\sqrt{R.P. + L.T.} - \sqrt{L.T.}) \sigma_{\text{month}} = 16.25$$

It is simple enough to generalize from the expression above. The reserve stock effect is negligible when:

- (1) K tends to zero—meaning that almost no protection is required against stock-outages.
- (2) The lead time is much greater than the review period.
- (3) The demand variance is very low—indicating a good deal of certainty as to what the demand will be in each period.

It is quite clear that the offsetting reserve stock cost may not be zero or negligible, but still less than the savings. In this circumstance, it is reasonable to use periodic ordering.

The decision as to the type of inventory policy which is to be used requires further analysis in terms of the developments above. To pursue this we will name three characteristics that play an important part in our systems analysis. These are capacity, load, and delay.

73. CAPACITY AND LOAD

The sum of the fixed and variable order costs—which is the total order cost—represents the data-processing *load*, expressed in dollars, for any given model. This expenditure of money is required for the operation of the inventory system. Thus, referring to our previous example, the Q -model load for the one item was \$34.00. For the same item handled by the P -model the load

was \$25.60. Perhaps it is easier to think in terms of more than one item. Therefore, let us presume that the company carries 1000 different items and that the item used in the example is close to the average. Then, the total *Q*-model load is \$34,000 per year and the *P*-model load is \$25,600 per year.

These figures tell us nothing whatever about the capacity of the company's clerical staff. The data-processing department might have a payroll of \$50,000. Or, it might in turn have a machine-time assignment equivalent to \$100,000 per year of the rental cost. The difference between the load and the capacity is idle time. Let us review for a moment the way in which the fixed and variable ordering costs were obtained. The job that had to be done was studied. The time required to do the job—composed of many individual elements—was measured. Normally, each type of job element has its own pay scale and so the individual times are multiplied by the appropriate pay scales and a total ordering cost is derived by summing across all of the individual time elements. The figure which we obtain in this manner is the dollar load. The labor load would be specified by adding together the number of processing hours required within each pay scale group. We could then derive either the total hour load or else a multiple category description showing the way in which the load was distributed. The fact is then that a company may not be in a position to take on the load which is indicated. On the other hand, the company may find that it is over-staffed or over-mechanized for whatever load is indicated. Before we continue to discuss the relationships of load, capacity, and idle time let us detail the meaning of these terms.

	MANUAL ROUTINE				
	1	2	3	4	...
Data-processing step	f_1	f_2	f_3	f_4	...
Expected frequency per week	h_1	h_2	h_3	h_4	...
Hours required per step	$h_1 f_1$	$h_2 f_2$	$h_3 f_3$	$h_4 f_4$	...
Total hours required per week	40	40	40	40	...
Hours available per worker per week	$h_1 f_1 / 40$	$h_2 f_2 / 40$	$h_3 f_3 / 40$	$h_4 f_4 / 40$	...
Number of workers required per week	P_A	P_A	P_B	P_C	...
Labor skill wage rate per hour					

All steps having the same labor skill rating and consequently the same wage scale can then be grouped together, yielding the following moe.

COMBINED LOAD	
Total hours required per week:	
Pay scale A:	$h_1 f_1 + h_2 f_2 = D$
Pay scale B:	$h_3 f_3 = E$
Pay scale C:	$h_4 f_4 = F$
Grand total:	$h_1 f_1 + h_2 f_2 + h_3 f_3 + h_4 f_4 = G$
Total number of workers required per week:	
Pay scale A:	$D / 40$
Pay scale B:	$E / 40$
Pay scale C:	$F / 40$
Grand total:	$G / 40$

Total dollar load:

$$\text{Pay scale A: } DP_A = H$$

$$\text{Pay scale B: } EP_B = I$$

$$\text{Pay scale C: } FP_C = J$$

$$\text{Grand total: } H + I + J = K$$

COMBINED CAPACITY

COMBINED IDLE FACTOR

Number of workers:

$$\text{Pay scale A: } W_A$$

$$\text{Pay scale B: } W_B$$

$$\text{Pay scale C: } W_C$$

$$\text{Grand total: } W_A + W_B + W_C$$

$$W_A - D/40$$

$$W_B - E/40$$

$$W_C - F/40$$

$$W_A + W_B + W_C - G/40$$

Number of worker hours:

$$\text{Pay scale A: } 40W_A$$

$$\text{Pay scale B: } 40W_B$$

$$\text{Pay scale C: } 40W_C$$

$$\text{Grand total: } 40(W_A + W_B + W_C)$$

$$40W_A - D$$

$$40W_B - E$$

$$40W_C - F$$

$$40(W_A + W_B + W_C) - G$$

Payroll per week:

$$\text{Pay scale A: } 40W_AP_A$$

$$\text{Pay scale B: } 40W_BP_B$$

$$\text{Pay scale C: } 40W_CP_C$$

$$\text{Grand total: } 40(W_AP_A + W_BP_B + W_CP_C)$$

$$40W_AP_A - H$$

$$40W_BP_B - I$$

$$40W_CP_C - J$$

$$40(W_AP_A + W_BP_B + W_CP_C) - K$$

As we would expect, the computer routine is similar. The only significantly different feature is that equipment is involved instead of workers. It should therefore be sufficient to merely indicate the basic structure of the computer routine. The translation from one system to another can be done rapidly.

COMPUTER ROUTINE

Data-processing step

Expected frequency per week

Hours required per step

Total hours required per week

Equipment class required

Equipment cost per hour

Total time assigned

COMBINED LOAD

Total hours required per week broken down by equipment classes. Total cost for time required per week broken down by equipment classes. These separate costs are summed for grand total cost. This is the dollar load factor.

COMBINED IDLE FACTOR

The difference between the capacity measures and the load measures. This is broken down by equipment classes.

COMBINED CAPACITY

Total hours available per week broken down by equipment classes. Total cost for available time per week broken down by equipment classes. These are summed for grand total cost. This is the dollar capacity factor.

Variants of the procedures above are certainly called for in particular applications. However, being loyal to the systems philosophy we have attempted to furnish a general description of the kind of detail which is required in an actual study.

74. UNDERLOADED DATA-PROCESSING SYSTEMS

The question now arises, given an expected load which is less than the fixed capacity of a data-processing system—or, more specifically, the ordering department—what effect does the cost of idle time have on the overall optimization of the system? Stated another way, how does underutilization of the data-processing system resources affect the inventory policies of our company? It is important here to note that capacity is assumed to be fixed and unchangeable. This is not an unusual condition. While the company has some control over the size of its ordering department, its reflexes are slow in adjusting to new requirements that cut down on the size of the system. Many explanations can be given but the most obvious one is the unpleasant aspect of wholesale firing. Company morale is affected. Community relations are affected, and if a union exists even more pointed arguments are raised. So the company will gradually absorb the idle capacity by shifting workers to other jobs as opportunities present themselves and, at the same time, normal turnover attrition does the rest of the job. Inflexible as this may seem, a computer-based system is even more rigid. The company invests in or rents a computer. After all possible assignments have been made on the computer there may be excess or idle time. The only way of removing this idle time is to offer computer time to other companies. This, of course, is a business in itself and many companies do not enjoy the prospect of becoming computer service companies. One alternative is to leave it idle, but, with respect to the inventory system, is this the best possible alternative? The answer is, generally, no.

Working with the same numbers which were used in our previous chapter, let us explore the importance of unutilized capacity. Our data-processing organization will be assumed fixed at \$50,000 per year including overhead costs. Since our previous load figures were \$34,000 and \$25,600, we have idle time costs of \$16,000 and \$24,400 for the *Q*-model and the *P*-model respectively. Now, we must work backwards for we want to learn how much, if anything, can be saved by fully utilizing the data-processing capacity. There are a number of advantages in assuming for this example that an average cost per hour exists to describe the pay of the clerical staff, each of whom will be said to work at the same average rate. The same kind of assumption will be applied to the equipment classes of the computer system. In this way we can avoid the multilabor-class or multiequipment-class derivation of load, capacity, and idle cost. Were we to introduce these, it would inhibit the clarity of our exposition and tend to obscure the main lines of our analysis. Therefore, we are either assuming that the \$50,000 per year is the payroll of a homogeneous clerical staff or that it represents machine-time assigned on a general purpose computer whose component operation times are roughly comparable.

We have already defined load, capacity, and cost of unutilized resources, namely,

Data-processing load = Fixed processing costs + Variable processing costs

and

$$\text{Data-processing capacity} = \text{Data-processing load} \\ + \text{Cost of idle facilities, (ID.)}$$

Thus,

$$\$50,000 = \frac{zG_1}{v} + \frac{z}{x} C_r + \text{ID.}$$

For the *Q*-model, ID. = \$16,000 and for the *P*-model, ID. = \$24,400. Our objective is to bring these idle costs to zero. To achieve this we have only one control which can be exercised and that is to decrease the size of x until the term zC_r/x has absorbed the idle cost. In this way we are assigning the idle cost to the variable cost component. Our control over the idle cost stems from the ability to order more frequently. Thus, by increasing the frequency of the variable operations, we succeed in using more and more of the unutilized capacity of the data-processing facilities. Our total cost equation, including fixed and variable ordering costs, carrying costs, reserve stock costs and idle costs is:

$$\text{Total cost} = \frac{zG_1}{v} + \frac{zC_r}{x} + \frac{cC_c x}{2} + cC_c K(\sigma) + \text{ID.}$$

We can quickly determine, for the average *Q*-model item with L.T. = $\frac{1}{2}$, $K = 3.25$, and σ month = 10 units:

Total cost (Capacity = 34)

$$= \frac{(1200)(.10)}{12} + \frac{(1200)(2)}{100} + \frac{4(.12)(100)}{2} + 4(.12)(3.25)(7) + 0 \\ = \$68.92$$

And for the *P*-model:

Total cost (Capacity = 25.60)

$$= \frac{1200(.01)}{12} + \frac{1200(2.10)}{102.5} + \frac{4(.12)(102.5)}{2} + 4(.12)(3.25)(12) + 0 \\ = \$68.92$$

This result merely restates our previous derivation when we balanced the *Q*-model against the *P*-model so as to be indifferent in our choice of systems. The idle cost was zero since the data-processing capacity was assumed exactly equal to the load. But now, we recalculate the total cost for the case where the capacity is larger than the load; i.e., fixed at \$50,000. Then, for the *Q*-model,

$$\text{Total cost (Capacity} = 50) = \$68.92 + \$16.00 = \$84.92$$

and for the *P*-model,

$$\text{Total cost (Capacity} = 50) = \$68.92 + \$24.40 = \$93.32$$

We are no longer indifferent in our choice of systems. For the 1000 item case, the cost of the *P*-model is \$8400 per year in excess of the cost of the *Q*-model; i.e., $(93.32 - 84.92)1000$. However, we must still investigate the effect of utilizing the idle time by increasing the frequency of the variable (controlled) processing functions. We then have for the *Q*-model:

$$50 = \frac{zG_1}{v} + \frac{zC_r}{x} = \frac{1200(.10)}{12} + \frac{1200(2.00)}{x}$$

whence,
$$x = \frac{1200(2.00)}{40} = 60 \text{ units}$$

and for the *P*-model:

$$50 = \frac{zG_1}{v} + \frac{zC_r}{x} = \frac{1200(.01)}{12} + \frac{1200(2.10)}{x}$$

Thus,
$$x = \frac{1200(2.10)}{49} = 51.4 \text{ units}$$

Substituting these new *x* values back into the total cost equations we obtain:

$$\begin{aligned} \text{Q-model: Total cost (Capacity} = 50) &= \frac{1200(.10)}{12} + \frac{1200(2)}{60} + \frac{4(.12)60}{2} \\ &+ 4(.12)(3.25)(7) = \$75.32 \end{aligned}$$

and for the *P*-model:

$$x = 51.4; \quad n = z/x = \frac{1200}{51.4}$$

and
$$\text{R.P.} = \frac{12}{n} = \frac{12(51.4)}{1200} = .514 \text{ months}$$

Therefore,

$$\sigma_{\text{R.P.} + \text{L.T.}} = \sqrt{\text{R.P.} + \text{L.T.}(\sigma_{\text{month}})} = \sqrt{.514 + .500(10)} \cong 10 \text{ units}$$

Then,

$$\begin{aligned} \text{Total cost (Capacity} = 50) &= \frac{1200(.01)}{12} + \frac{1200(2.10)}{51.4} \\ &+ \frac{4(.12)(51.4)}{2} + 4(.12)(3.25)(10) = \$78.14 \end{aligned}$$

Our result plainly demonstrates that when part of the data-processing capacity goes unutilized, savings can be obtained by moving away from the optimal order quantity solution. We tabulate the results below:

TOTAL COST	Q-MODEL	P-MODEL	ORDER QUANTITY
(1) Complete utilization: Load = Capacity	\$68.92	\$68.92	Optimal
(2) Underutilization Load < Capacity	84.92	93.32	Optimal
(3) Complete utilization: The load is increased to absorb idle capacity.	75.32	78.14	Non-optimal—decrease order size; increase order frequency
(4) Overload: Load > Capacity	88.52*	72.93*	Non-optimal—increase order size; decrease order frequency

* See the section which follows for both the derivation and discussion of this row.

Recapitulating, we started out with a situation in which we were indifferent to the choice of *P*- or *Q*-models (1). But then we recognized the fact that idle capacity was going unnoticed. Once we included this additional cost, the *Q*-model was preferred (2). We then absorbed the idle cost. The *Q*-model still retained its advantage, but a substantial saving was made for both models by operating without the optimal lot size criterion (3). This result should not, however, be construed to signify that the full utilization of capacity is a desirable goal in itself. Other costs, which we have not yet discussed, become influential when the extent of a system's utilization approaches 100%. These costs reflect an inability to handle fluctuations and growth in demands.

75. OVERLOADED DATA-PROCESSING SYSTEMS

As long as the actual demands made on the data-processing facility are constant and unchanging, the full utilization of capacity is in order. But few companies would look with favor on the prospects of experiencing absolutely no growth in demand over a long period of time. If the processing capacity can be augmented as required, then the company can increase capacity as the load grows. Such adjustments of capacity are never easy to accomplish. It is likely to be more difficult to achieve with a computer-based system than with a predominantly clerical system. If immediate adaptability is not feasible then planning ahead to anticipate the company's future needs is another possibility. However, good long-range plans based on accurate long-range forecasting would be required. Seldom can reliance be placed on predictions of long-term capacity needs. Too many factors intervene and the longer the time period the more unexpected the results can be. Nevertheless, it is clear that there is a need for some unused capacity to cover the growth of data-processing requirements. Let us investigate the effect of not having sufficient capacity to handle the load. Again, returning to our

previous example, we will suppose that the capacity of the data-processing system is exactly \$20,000. Then, for the Q -model:

$$20 = \frac{zG_1}{v} + \frac{zC_r}{x} = \frac{1200(.10)}{12} + \frac{1200(2.00)}{x}$$

and
$$x = \frac{1200}{10} (2) = 240 \text{ units}$$

So the total cost would be:

$$\begin{aligned} \text{T.C.} &= \frac{zG_1}{v} + \frac{cC_c x}{2} + cC_e K(\sigma) + \text{ID.} \\ &= \frac{1200(.10)}{12} + \frac{1200(2)}{240} + \frac{4(.12)240}{2} + 4(.12)(3.25)7 \\ &= \$88.52 \end{aligned}$$

The cost has increased to the largest value we have obtained for the Q -model so far. The perpetual inventory system continues to operate as it did before. The fixed cost component is unchanged and so, in order to satisfy the \$20,000 budget, the order quantity has been increased substantially, thereby reducing the number of purchase orders. For the 1000 item company an increase in payroll or machine-time of \$14,000 per year will produce a net saving of \$19,600 per year. Now for the P -model:

$$20 = \frac{zG_1}{v} + \frac{zC_r}{x} = \frac{1200(.01)}{12} + \frac{1200(2.10)}{x}$$

and
$$x = \frac{1200(2.10)}{19} \approx 133 \text{ units}$$

Next, we determine the review period:

$$\text{R.P.} = \frac{12}{n} = \frac{12x}{z} = \frac{(133)(12)}{1200} = 1.33 \text{ months}$$

Therefore:

$$\sigma_{\text{R.P.} + \text{L.T.}} = \sqrt{\text{R.P.} + \text{L.T.}} [\sigma_{\text{month}}] = \sqrt{1.83}(10) = 13.5 \text{ units}$$

and

$$\begin{aligned} \text{T.C.} &= \frac{1200(.01)}{12} + \frac{1200(2.10)}{133} + \frac{4(.12)(133)}{2} + 4(.12)(3.25)(13.5) \\ &= \$72.93 \end{aligned}$$

This cost, we may note, is not unreasonably high in comparison to the other costs that we previously derived. Since the P -model has a smaller fixed cost component, it is more flexible than the Q -model when the available capacity is exceeded. A new schedule is set up with a somewhat longer review period to compensate for the lack of data-processing facilities. Of course

there is a cost for unanticipated demand. The Q -model is only more vulnerable than the P -model in this regard. For the 1000 item company, an increase in payroll or machine-time of \$5600 per year will produce a net saving of \$4010 per year.

76. FLUCTUATIONS IN LOAD

Up to this time we have been talking about capacity and load, etc., in dollar terms. This was convenient and satisfactory for our previous needs. But the question arises, "What effect do fluctuations in demand have on systems that operate at different percent utilizations of capacity?" To properly discuss this we can no longer restrict our analyses to dollar terms. We must begin to consider the rates at which demands are made and the rates with which such demands can be serviced.

Let us first consider operations which are independent of the order frequency, for example, issue transactions as processed by the Q -model. We have z/v or $1200/12 = 100$ issue transactions expected per year for our average item. This is true no matter what the order frequency. There would be 100,000 issue transactions for 1000 items. Referring to our previous load and capacity charts we could call the treatment of the issue transactions, data-processing step 1. The expected frequency per week is

$$f_1 = \frac{100,000}{52} = 1923$$

The hours required per transaction might have been $h_1 = .08$. Thus the total number of hours required per week would be $h_1 f_1 = 154$. Assume that four clerks are assigned to the task. Together, they provide an available supply of 160 hours per week. This leaves an excess capacity of about six hours. By our definition, this is idle capacity or unutilized data-processing resources. Should we then get rid of this excess? We can do that by assigning the fourth clerk to part-time operation at another task. The answer to our question can only be supplied by examining the arrival rates and the servicing rates of the issue transactions.

If arrivals were evenly spaced, a new one would appear every $40/1923 = .021$ hours = 1.25 minutes. Another way of looking at this is to say that $1923/40 \approx 48$ evenly spaced issue transactions are expected to arrive every hour. Now, it is evident that only the most unusual circumstances would foster such regularity. Some hours, 50 transactions may arrive—other hours may find 45 transactions coming in for processing. But we can say that there is an expected arrival rate of 48 transactions per hour and then specify the variance of the arrival distribution. The same reasoning applies to servicing rates. The expected rate is .08 hours per transaction and conversely, $1/.08 = 12.5$ transactions per hour. Following the accustomed notation of queuing theory we will name the arrival and service parameters, λ and μ , respectively. We have, $\lambda = 48$ transactions per hour and $\mu = 12.5$ transactions per hour. We let $M = 4$ clerks, each of them capable of providing

service at the μ rate. This gives us ρ , which we have previously identified as the utilization factor of our system. For our example,

$$\rho = \frac{\lambda}{M\mu} = \frac{48}{(4)(12.5)} = .96$$

In queuing theory terminology, we categorize the situation as one having the possibility of an infinite queue and with multiple channels. The queue length can be infinite because there is no limit within reason to the number of

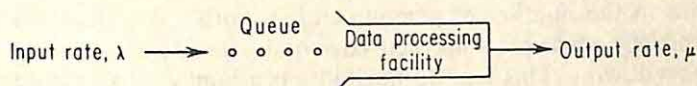


FIG. 9-1 Single-channel data-processing system. (Infinite queue length is possible.)

demands which can be made upon the system. Practically speaking, we know that an infinite queue would never form. On the other hand, to state that there are only a finite number of issue transactions which could (not would) wait in line for servicing is not accurate. The problem has four channels as we have stated it. A single channel problem would have only one clerk assigned to the job of clearing the issue transactions. Figures 9-1 and 9-2 schematically represent a single and multiple channels system, respectively. Having categorized our problem, we can now find the appropriate formulas.

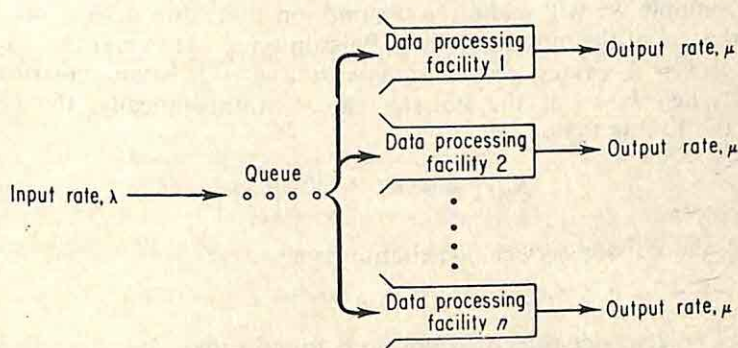


FIG. 9-2 Multiple-channel data-processing system. (Infinite queue length is possible. Arrival transaction goes to first free data-facility.)

References to queuing literature which present many such formulas can be found in the bibliography of this book.

Before we proceed to the analytical resolution of our problem, it will be illuminating to discuss the forms which the arrival and service distributions can possess. For the application of queuing theory a specific family of distributions is used. It is called the Erlang family. The basic building block of the Erlang family is the exponential distribution. It has the important attribute, which we have had occasion to mention before, that successive

events are independent of all preceding events. This means complete randomness of arrivals. In this case, the number of arrivals per unit time will be assumed to follow the Poisson distribution. Similarly, for the service distribution we have complete randomness of servicing times. The distribution of servicing times will also follow the Poisson distribution.

The latter assumption (i.e., exponential service) would not be true for a computer since its service times could be almost constant—within any meaningful precision limits. It also might not be true for the clerical staff. Clerks can be quite flexible in adjusting their work rates to accommodate small fluctuations in the number of arriving transactions. When 50 transactions are received, they can speed up their pace and when 45 are received they will tend to slow down. This kind of flexibility is a human attribute and not one shared by a computer. On the other hand, the expected variance of the human operator, even when he is fed work at a constant rate, is likely to be quite large. The Poisson distribution will frequently provide an excellent description. If the distribution of servicing times is not of the pure Poisson type, then it is very likely that another member of the Erlang family will provide a sufficiently good description. Actually, the various members of the Erlang family represent varying degrees of randomness. The exponential case is categorized as pure random behavior, but what is really meant is that the mean and variance of the distribution are equal. This, of course, is the Poisson attribute. Distributions which have larger variances than the mean are called hyper-exponential. As the variance tends toward zero the constant service-time, such as the computer servicing rate, is obtained. For our example we will make the assumption that both arrival rates and service rates are of the pure random or Poisson type. However, we will first compare a $k = 2$ Erlang distribution with a $k = 1$ Erlang distribution. The latter, i.e., $k = 1$ is the Poisson case. Mathematically, the general form for the Erlang family is:

$$S_0(t) = e^{-k\mu t} \sum_{n=0}^{k-1} \frac{(k\mu t)^n}{n!}$$

Therefore, the Poisson service distribution is given by:

$$S_0(t) = e^{-\mu t}$$

while the $k = 2$ service-time distribution is found to be:

$$S_0(t) = e^{-2\mu t}(1 + 2\mu t)$$

With respect to the above we must note that $S_0(t)$ is a cumulative probability distribution which expresses the probability that the service-time interval is greater than t .

Figure 9-3 pictures the service-time distribution $k = 2$, and compares it with the service-time distribution, $k = 1$, which is the pure random case. It will be noted that the Erlang 2 lies above the Erlang 1 until a crosspoint is reached. This indicates that less variability and less randomness exist in the Erlang 2 case. There is a higher probability for Erlang 2 than for Erlang 1 that the service time is greater than the low values of t . There is also a lesser

probability that the service-time is greater than the higher values of t . As we tend toward the limiting case, the higher Erlang numbers move more and more of the area under the curve into the region to the left of the expected value T_s , where $T_s = 1/\mu$. Ultimately we would have a rectangle bounded by $t = 0$, $t = T_s$ and $S_0(t) = 0$, $S_0(t) = 1.00$. This represents the case of constant service-time equal in value to T_s . It should also be noted that the area bounded by the rectangle is equal in value to $(1 \times T_s) = T_s$. This is not

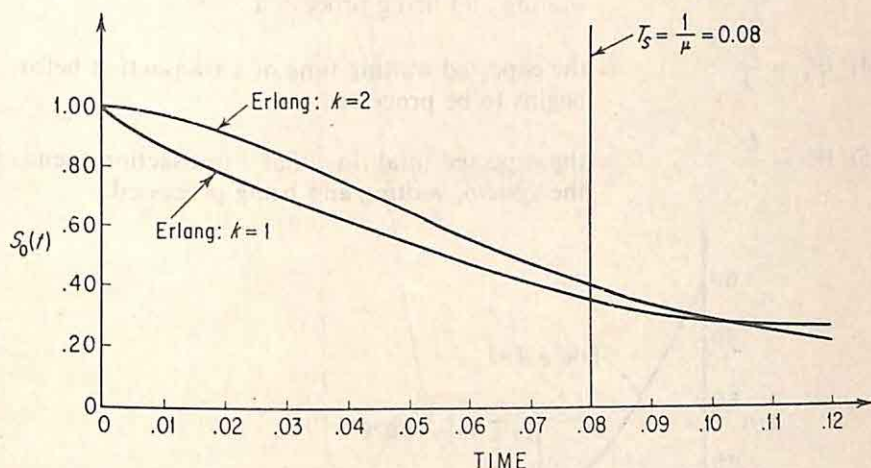


FIG. 9-3 Issue transaction cumulative service-time distribution. ($S_0(t)$ = probability that service time is greater than t .)

peculiar to the constant service-time situation. The area under all of these curves is equal to the expected service time.

Figure 9-4 shows the cumulative arrival-interval distribution which is of the Poisson type. As before,

$$A_0(t) = e^{-\lambda t} \quad (k = 1)$$

where $A_0(t)$ is the cumulative probability distribution which describes the probability that the interval between arrivals is greater than t . Here too, the area under the curve is equal to T_a , where $T_a = 1/\lambda$ and is therefore the expected value of the interval between arrivals. Each of the cumulative distributions can be restored to its equivalent probability density distribution. Figure 9-5 illustrates these density distributions, which can be obtained from the cumulative distributions by subtracting successive values from each other.

Having prepared some background for this analysis of our data-processing system, we are ready to answer the question which we previously raised; namely, should four clerks be employed to handle issue transactions or, more generally, at what percent of capacity should we operate our system?

The relevant queuing equations are (for ρ approaching 1):

- (1) $P_0 = \frac{(1 - \rho)M!}{M^M}$ = the probability that no transactions are in the system, either waiting or being processed.
- (2) $L_q = \frac{\rho}{(1 - \rho)}$ = the expected number of transactions which are waiting to be processed.
- (3) $L = L_q + \rho M$ = the expected number of transactions which are waiting and being processed.
- (4) $W_q = \frac{L_q}{\lambda}$ = the expected waiting time of a transaction before it begins to be processed.
- (5) $W = \frac{L}{\lambda}$ = the expected total time that a transaction spends in the system, waiting and being processed.

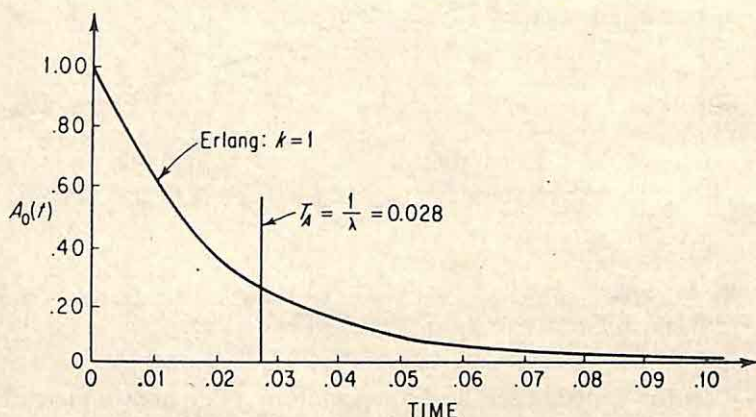


FIG. 9-4 Issue transaction arrival-interval distribution. ($A_0(t)$ = probability that an interval between arrivals is greater than t .)

First, we solve these equations with $M = 4$ clerks and then with the $M = 5$ clerks. It should be noticed that when $\rho = 1.00$ the system is fully utilized. When $\rho > 1.00$, the load is greater than capacity and when $\rho < 1$, the system's capacity is greater than the expected load. For $\rho > 1.00$ the waiting line will grow indefinitely, and our equations do not describe this condition; e.g., with $M = 3$, we have $\rho = 1.20$. It serves clear warning that to overload the fixed processing elements is absurd. Even when we operate the system at almost full utilization we are risking catastrophe unless our decision to do so is based upon a comprehensive analysis. There is always the risk of underestimating the actual load even by a small amount—with the prospect in store of requiring emergency measures in order to regain control over the processing system.

The table below tells us that even when our estimates are accurate and the system is operating at underload conditions a waiting line will form. With a

system utilization of 96%, which we achieve with four clerks, we find that each transaction will, on the average, wait untouched for one-half hour.

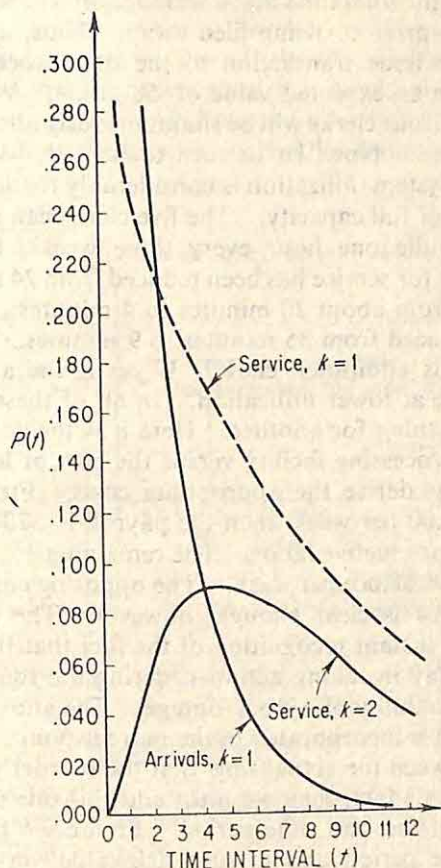


FIG. 9-5 Issue transaction density distributions.

$P(t)$ = probability that an interval of time equal to t will elapse between arrivals or between the beginning and completion of a servicing operation.

NUMBER OF CHANNELS		$M = 4$	$M = 5$	$M = 6^*$
ρ	(Utilization factor)	.960	.768	.640
P_0	(Probability of total idleness)	.004	.008	.019
L_q	(Transactions waiting)	24	3.3	.04
L	(Transactions in system)	27.8	7.1	3.88
W_q	(Hours waiting)	.50	.07	.0008
W	(Hours in system)	.58	.15	.08

* These measures had to be obtained in a different way since ρ is not close to 1.00, a condition our formulas required. The reference given here presents formulas which will handle this case. In addition, tables of values are given which are quite simple to use. See Philip M. Morse, *Queues, Inventories and Maintenance* (New York: John Wiley & Sons, Inc., 1958).

This does not appear to be an unreasonable waiting time. Since the time to process a transaction is .08 hours, we add this amount to the waiting time in order to estimate the total time that a transaction will be expected to remain in the system; i.e., prior to being filed away. Thus, assuming immediate transmission of the issue transaction to the data-processors, the delay in processing will have an expected value of .58 hours. We also observe that .004% of the time all four clerks will be simultaneously idle. This is about one hour every six weeks. Now, let us turn to $M = 5$. When an additional clerk is added, the system utilization is considerably reduced. It is operating at only about 75% of full capacity. The five clerks can now be expected to be simultaneously idle one hour every three weeks, but the number of transactions waiting for service has been reduced from 24 to 3.3. The waiting time has been cut from about 30 minutes to 4 minutes. The expected time in the system is reduced from 35 minutes to 9 minutes.

Do we want this additional clerk? If yes is the answer, then we are prepared to operate at lower utilization. In all of these capacity problems we are trading one thing for another. Here it is the cost of additional idle time for the data-processing facility versus the cost of less waiting time per transaction. Let us derive the appropriate costs. First, when $M = 4$, if clerks are paid \$50.00 per week, then the payroll is \$200.00. Only 96% of the payroll goes to productive labor. The remaining 4% is an idle-time cost. Thus $(.04 \times \$200 = \$8.00)$ per week. The opposing costs are more difficult to define. The issue is clear enough, however. The perpetual inventory system requires the instant recognition of the fact that the reorder point has been reached. Delay in taking action—ordering the replenishment quantity x —increases the probability of a stock-outage. The allowance for demand in the lead time period is incorporated in the reorder point. Thus, if a delay of one day occurs between the actual time that the reorder point is reached and the recognition of this fact, then we must add this one day's demand to the expected demand in the lead time period. Figure 9-6 illustrates how delay shifts the lead time period and in turn delays delivery with the resultant effect that the probability of running out of stock increases.

This is the effect of delay. How can we determine its cost? The expected time in the system—waiting for service and being processed—for each issue transaction is .58 hours. This includes all those transactions which trigger reorder point ordering activity. For those transactions which do not reach the order point, the delay is of no significance. But for those which do, an additional amount of stock must be carried. Then, the expected unit demand in the additional .58 hour delay period would be $(.58)(100/160) = .3625$ units, where 100 units are demanded per month, and we assume 160 hours per month. This reduces the effective size of the reserve stock by a very small amount, namely .3625 units. If this amount is added to the reserve stock the additional carrying cost per week would be:

$$c \left(\frac{C_c}{52} \right) (.3125) = \frac{4(.12)(.3625)}{52} = \$0.00335$$

This is a very low cost, even when multiplied by 1000 to account for 1000

hypothetical items carried by the hypothetical firm. This total cost of \$3.35 can then be added to the idle-time cost of the clerks namely, \$8.00 per week. The grand total is \$11.35 and clearly indicates that a fifth clerk would not be a useful addition in this department. With five clerks the idle-time cost will be \$58.00 per week. Therefore, even if the transaction delay costs went to zero with five clerks, the four-clerk total cost would always be preferred.

Let us see, however, what would happen if we had specialization of work-loads. That is, each of the four clerks handles transactions for a particular

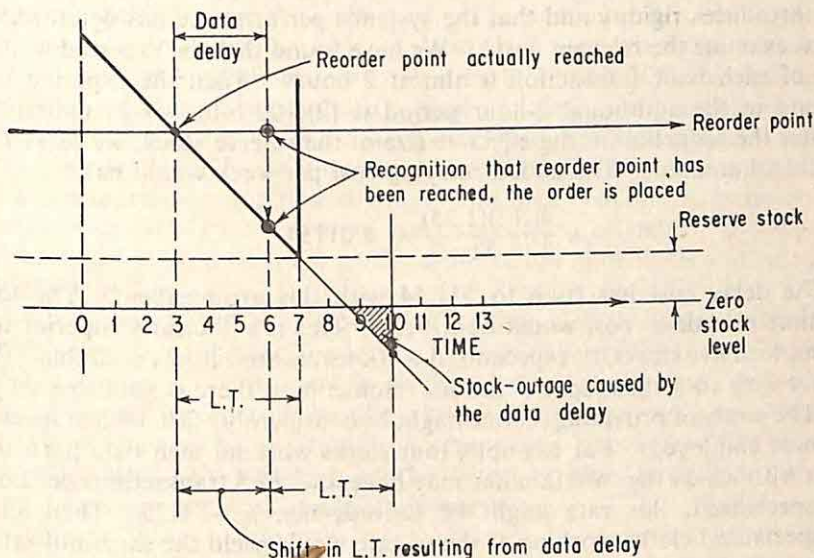


FIG. 9-6 The effect of delay on the probability of running out of stock. (Normally, delivery would be made at $t = 7$. The delayed delivery occurs at $t = 10$.)

set of the company's total line. We will assume that each clerk carries $\frac{1}{4}$ of the load. Then our queuing problem becomes one with infinite queues but single channels. The arrival rate is quartered; $\lambda = 48/4 = 12$ transactions per hour. The servicing rate remains the same; i.e., 12.5 transactions per hour. We now derive:

$$\rho = \frac{12}{12.5} = .96$$

which is the same utilization factor as before. Certainly it must be. But we now have different queuing equations since an arrival cannot be shifted between clerks. An arrival of type A must wait for clerk A . The equations we require are:

$$P_0 = (1 - \rho) = .04$$

$$L_q = \frac{\rho^2}{(1 - \rho)} = 23.04 \text{ transactions}$$

$$L = \frac{\rho}{(1 - \rho)} = 24.00 \text{ transactions}$$

$$W_q = \frac{L_q}{\lambda} = 1.92 \text{ hours}$$

$$W = \frac{L}{\lambda} = 2.00 \text{ hours}$$

The notation is the same as that used previously; it is apparent that specialization introduces rigidity and that the system's performance has deteriorated. Let us examine the relevant costs. We have found that the expected waiting time of each issue transaction is almost 2 hours. Then the expected unit demand in the additional 2-hour period is $(2)(100/160) = 1.25$ units. To counter the reduction in the effective size of the reserve stock, we carry this additional amount. The added carrying cost per week would be:

$$\frac{4(.12)(1.25)}{52} = \$0.1154$$

The delay cost has risen to \$11.54 with this arrangement. The total idle-time plus delay cost would be \$19.54. This is still clearly superior to a system with five clerks, but specialized work teams appear less desirable. We do not wish to dismiss specialization. Sometimes, there is good reason for it. The μ rate of providing service might be considerably faster when specialization is employed. For example, four clerks working their own particular items with which they are familiar may have $\mu = 12.5$ transactions per hour. Nonspecialized, this rate might be halved; i.e., $\mu' = 6.25$. Then, eight nonspecialized clerks working at the μ' rate would yield the same utilization factor and there would be additional delay costs to boot.

Now one more point should be made. The balance we have been investigating is noticeably sensitive to many factors including price. Thus, consider items which cost \$400.00 instead of \$4.00. For the unspecialized four-clerk system we would have found that the additional reserve stock cost was \$.335 per week. The 1000 item total would have been \$335.00 per week. This amount plus the \$8.00 idle-time cost would be \$343.00 giving a clear mandate to investigate the advantage of extra clerks. With five clerks, the cost would become:

$$(.15) \left(\frac{100}{160} \right) = .09375 \text{ units}$$

and

$$(400)(.12) \left(\frac{.09375}{52} \right) = \$0.8654$$

For 1000 items this is \$86.54. Adding the clerical idle-time cost associated with five workers we have $\$58.00 + \$86.54 = \$144.54$ per week. Thus, here five clerks have a lower total cost than four clerks. How about six clerks? Our previous table lists the required moe. Idle time cost would be $(\$300 \times .36) = \$108.00$. Delay costs would be \$46.15 which yields a

total cost of \$154.15. Therefore, five clerks is the closest integer-clerk solution.

Now, let us briefly consider the problem of those data-processing functions which are dependent upon the order quantity or the review period. We will think in terms of the *P*-model. The input is now constant. There are exactly n items to be reviewed—no more and no less. The only variable factor is the number of transactions which have accumulated between reviews and which must be added together. These are a minor portion of the total processing time. When the *P*-model is processed by a computer almost all processing delays* have been removed. The computer furnishes constant servicing times to an input which is essentially a nonfluctuating, nonvarying load. Thus, this kind of delay can only arise as a result of fluctuations in the number of transactions to be added together—and as we have pointed out, the very essence of the *P*-model is that such operations use up a negligible percent of the system's capacity. Consequently, with an almost constant input load and a constant processing time derived by means of computer operation—the fixed reorder cycle model permits exact scheduling of the use of facilities. An additional boon is the close to 100% utilization of facilities and an end to bottleneck problems.

Certain operations are prone to bottlenecks. Bottlenecks create a waiting line at one point in the system. They may very well create idle-time at other locations. Thus, with a perpetual inventory system, part of the purchasing department may be idle while issue transactions join a waiting line at the data-processing center. An important question in our systems analysis is whether the flow through the system becomes more or less erratic as the data units advance through the system.

77. REGULATION OF DATA FLOW

While analytical methods will sometimes resolve the question we have just raised, simulation will always provide us with a usable answer. Furthermore, if a sufficient sample is taken, simulation can frequently be more accurate than the mathematical approach. Our queuing models were flexible to the extent that the Erlang family could describe with sufficient precision the prevailing situation. But the fit may be too far off to be used. Some distributions simply defy description in Erlang terms. This is no deterrent to simulation procedures. Another reason exists for the increased accuracy that can be obtained with simulation. Many theoretical distributions have infinity as an upper bound. Although the probability of occurrences in the region of infinity are very small, nevertheless, the theoretical possibility influences our mathematical results. In reality, events in the vicinity of infinity would not be tolerated and would never occur. Therefore, actual distributions tend to be truncated at their upper tails. On the other hand,

* This does not mean that transmission delays have been removed. That is, the necessary information may not reach the right place in time. We will consider such types of delays later.

the mathematical description of a system's behavior is sometimes a better representation of the true situation that prevails when accurate observation is costly, time consuming, difficult, or impossible. Certainly, each approach has its advantages and its pitfalls which the systems analyst must weigh in choosing his course of action.

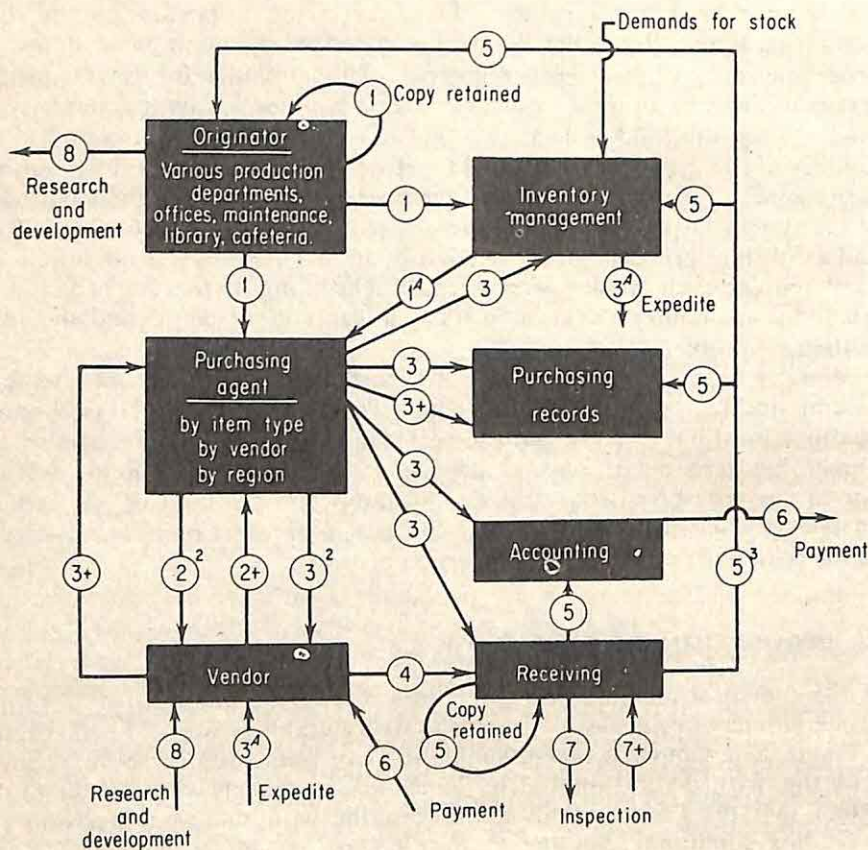


FIG. 9-7(a) A detailed block flow diagram of materials management.

The flow diagram pictured in Fig. 9-7 is too complex to analyze mathematically. By simulation techniques such an analysis is possible. We can write mathematical and statistical descriptions that relate the various flows in the network. But, in order to do so we must become more detailed. Figure 9-8 makes a good start in this direction. Here we have the various processing steps revealed. By studying such a flow chart it is possible to discover better ways of arranging sequences. It is also possible to associate each operation with a cost per unit time and with the distribution of times required to complete the operation. Abbreviations and symbols which

appear on the diagram are also explained there. Figure 9-8 has expanded a small portion of Fig. 9-7. It describes only those flows which are concerned with a part of the processing routine for receipt transactions. Similar flow diagrams can be drawn for all of the other segments of the network.

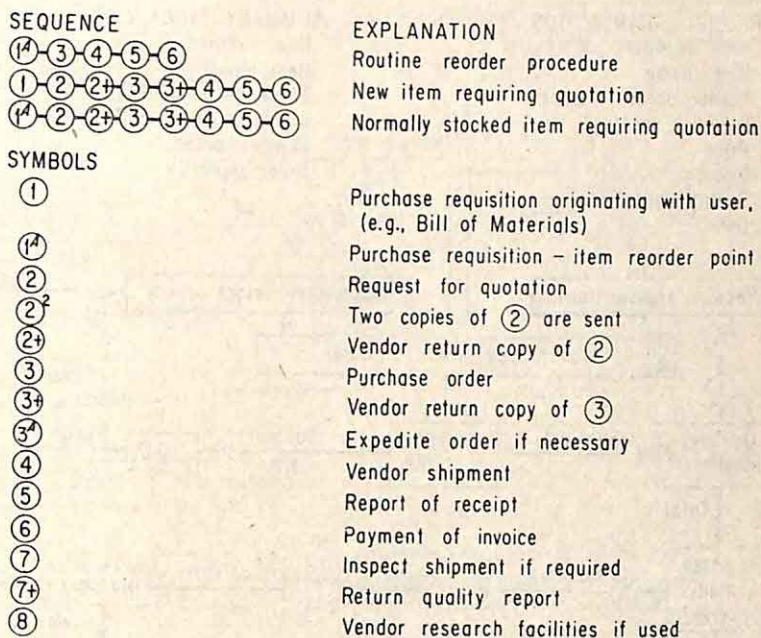


FIG. 9-7(b).

We have included Fig. 9-9 to present the reader with an example of an even more detailed flow diagram. This general type of flow chart is employed when computer systems are to be designed to handle the data-process flow. It is certainly not our intention to instruct the reader in computer flow-charting, coding, and programming. Nevertheless, some acquaintance with this ultimate degree of detail is useful background for the understanding of how basic the elements must be when the systems analysis is carried into the computer stage. Simulations by computer also require these excursions into the heartland of detail. Here, we have indicated the nature of an element by the shape of the box which surrounds it. The enclosure for S.O.H. + S.O.O. represents a magnetic tape storage. The receipt transactions are identified by their design as being documents. Quality adjustments are obtained from punched cards—a fact indicated by the sawed-off upper left-hand corner. Rectangular boxes usually signify operations of one kind or another while ovals are used to show that a decision must be made.* It will be noted that a

* A variety of flow-chart symbols are utilized. Several major computer companies have advanced their own systems. Perhaps standardization would be helpful but then, so would Esperanto.

configurations operate in both manual and computer systems. Of course, many of the decision junctures and loops operate inside the cerebrum of the operator. But other complex flow networks which do exist externally can be traced by systematic flow-charting. The computer must have the most

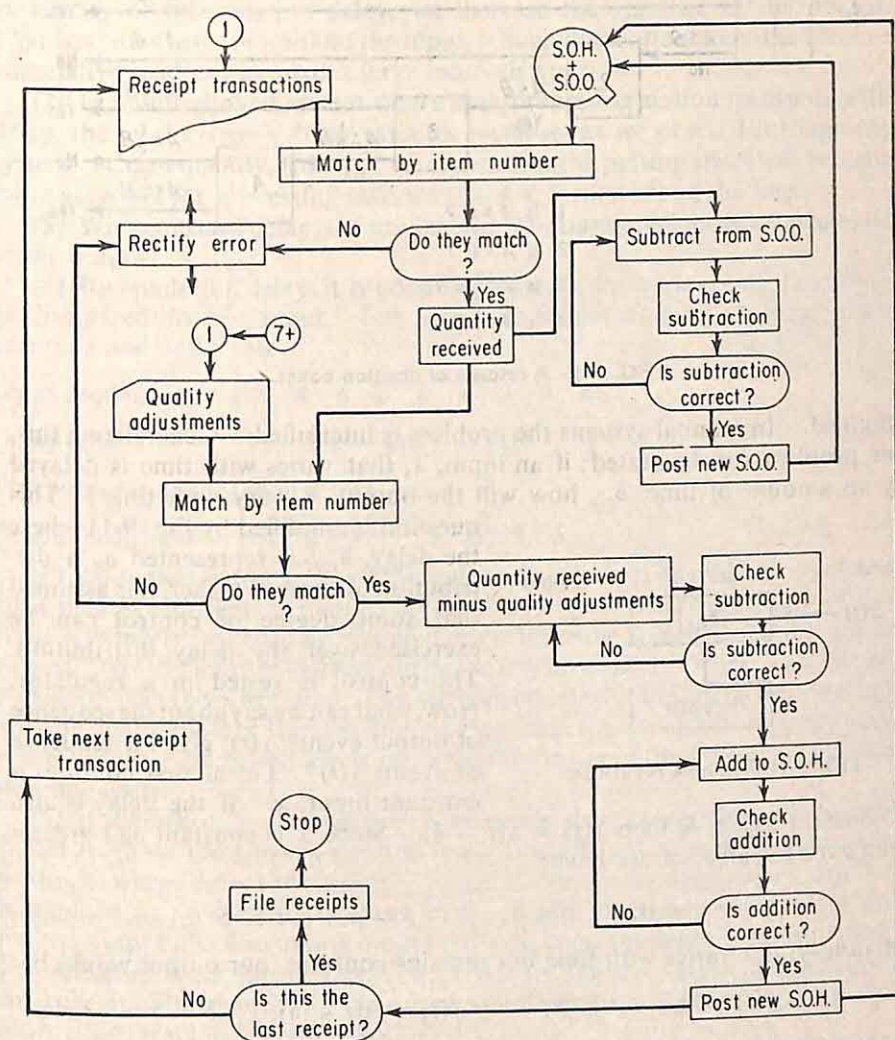


FIG. 9-9 Detailed computer flow diagram of receipt transactions.

precise instructions so that all eventualities can be handled. The business computer is very literal which explains its reliability as well as the need for patient systems preparation for all programs.

Operations and transmission times within the computer are well-regulated. Underlying good programming is sound structure. This in turn provides for

the sequencing of operations which can produce a result in minimum time. However, even a computer system has human operators intervening whose actions introduce variability with respect to the time that the output is

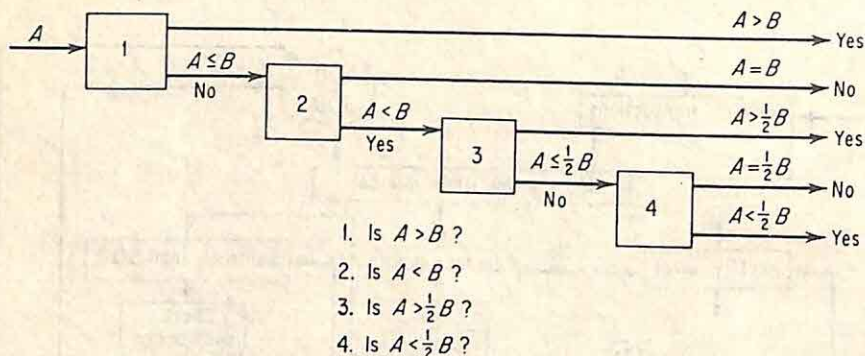


FIG. 9-10 A cascade of decision boxes.

obtained. In manual systems the problem is intensified. Generalizing this, our problem can be stated: if an input, x , that varies with time is delayed by an amount of time, $\delta_{x,y}$ how will the output, y , vary with time? This

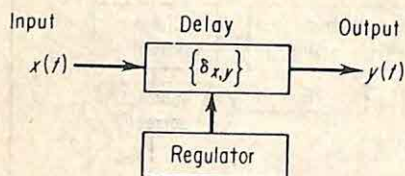


FIG. 9-11 A simple regulator.

question is amplified by Fig. 9-11 where the delay, $\delta_{x,y}$, is represented as a distribution of times. Further, it is assumed that some degree of control can be exercised over the delay distribution. The control is vested in a regulator. Now, what can we say about the sequence of output events, $y(t)$, given a sequence of events $x(t)$? Let us first consider a constant input, x . If the delay is also

a constant value, δ , then $y(t) = x(t - \delta)$. Since x is constant and greater than zero for all $t > 0$, we have;

$$y = 0; \quad t \leq \delta; \quad y = x, \quad t > \delta$$

Obviously, if x varies with time but remains constant, our output would be:

$$y(t) = 0, \quad t \leq \delta; \quad y(t) = x(t - \delta), \quad t > \delta$$

As soon as δ is no longer a constant but a probability distribution, the issue becomes more involved. Some mathematical conclusions can be drawn and some logical deductions are possible. A number of fields contribute substantially to the body of analytic knowledge which exists to deal with such problems. Foremost are the electronic control systems field, queuing theory and cybernetics. The last is a study of control systems under a wider range of possible conditions than are ordinarily included by the first two.

Our simulation procedures are by far the most flexible but they lack the ability to provide the generalizations that are so vital to understanding and progress. Consequently, let us first state a few general rules which should help us to interpret the simulation results:

(1) A constant input is one with zero variance. As we increase the variability of transmission delay, we increase the variance of the output. That is, it fluctuates more than the input. Such fluctuation takes the form of cumulative load pulses spaced large intervals apart.

(2) In an interlinked system where each processing station transmits with delay, the total system's delay variance increases as we proceed through the system. Consequently, from (1), we expect that the pulsing effect will become more accentuated for processing stations that are further along the line.

(3) With constant delay, the original input variance will remain unchanged at the output.

(4) By regulating delay, it is possible to reduce the variance of the output as compared to the input. For example, consider the following input sequence and delay rule:

Input sequence: 2 4 6 8 6 4 2 4 6 ...

Delay rule: $y(t) = \frac{1}{2}x(t) + \frac{1}{2}x(t - 3)$

The output sequence would then be:

Output sequence: 1 2 3 5 5 5 5 5 ...

(5) To reduce the variance of the output distribution an efficient regulator must possess sufficient variability to be able to cope with the variety of the input.* Only in this way can the regulator achieve the objective of a smoothed output. We desire a smoothed output so that a constant work load can be obtained. On the other hand, if we are concerned with the age of information when it finally reaches the decision-making stages, we may accept variability in the output. Increased levels of variability in the output produce greater idle-time costs.

(6) If a work-pool arrangement is available then processing time can be shifted to those locations where the load is heavy and removed from those locations where idle time exists. In this way, processing delay can be maintained at an almost constant level. It will be remembered from (3) that the output of a fluctuating input load with constant delay in transmission will reflect no reduction in variability. But in this case, efficient regulation can substantially reduce idle-time and simultaneously decrease the age of information reaching decision-points.

Let us now take an example and simulate the results for an interlinked system of flow. Our example will be relatively simple but the approach indicated will be applicable to a wide variety of situations. Of course, *our interest now is precisely how the flows through a network affect the results produced by an inventory model*. Our example is simple but at least somewhat

* As Ashby states, "Only variety can destroy variety," W. R. Ashby, *An Introduction to Cybernetics* (New York: John Wiley & Sons, Inc., 1956), p. 207.

realistic. We include in our system the use of a Q -model with fixed reorder point. We have a central warehouse and one branch warehouse. The data-processing center communicates with the central warehouse by means of a telemetering link. The example includes both materials and data flows. Because of space limitations we will not attempt to compare structural variants of the system, but it is suggested that the reader try a few different arrangements. One can assume that delays at each processing point can be changed by reallocation of a fixed budget. The use of a work-pool arrangement can be tested—recognizing the fact that only a centralized data-processing organization could avail itself to the fullest of the benefits of such an arrangement.

Figure 9-12 illustrates the system we will simulate. To avoid having to present an excessive number of numerical tables, we will use the direct routing, $B-E$, instead of the alternative routing (dotted line) $B-C-D-E$. The simulation of this latter routine is suggested as an exercise in the problem section of this chapter. Most of the information which is required for the simulation is presented as part of Fig. 9-12.

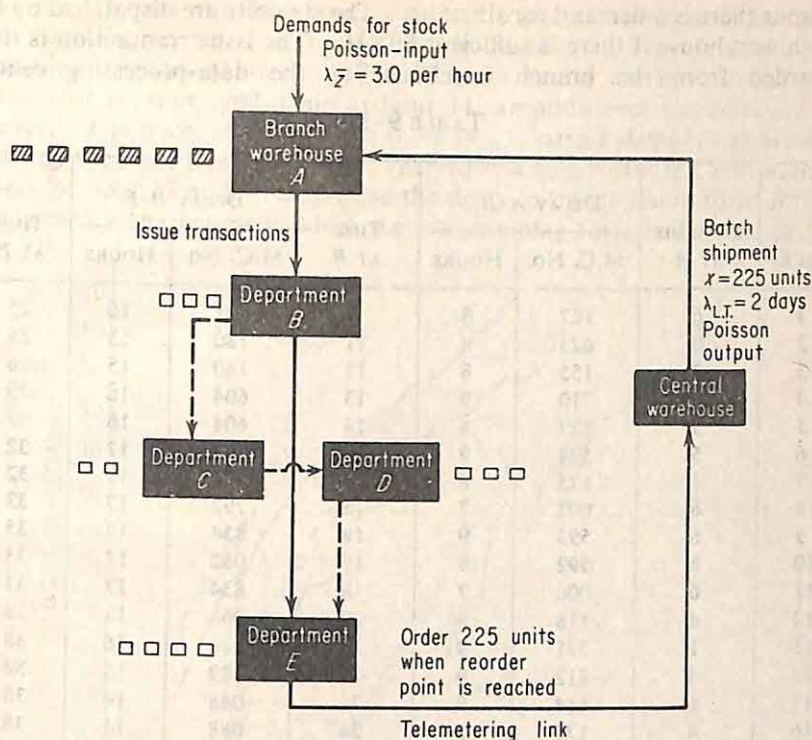
Our first step is to write down the input-output flow behavior in a form suitable for Monte Carlo procedures. It must be remembered that outputs from one element become the input to the next in line. Our example is limited to a single item.

DEMAND INPUT (UNITS PER HOUR)		ASSIGNMENT		MONTE CARLO NUMBERS AND ASSIGNED DEMAND							
0		000-049	962	6	950	6	086	1	077	1	
1		050-198	035	0	773	4	679	4	075	1	
2		199-422	221	2	119	1	621	3	220	2	
3		423-646	637	3	706	4	458	3	208	2	
4		647-814	550	3	927	6	011	0	478	3	
5		815-915	819	5	788	4	292	2	767	4	
6		916-965	063	1	799	4	198	1	264	2	
7		966-987	923	6	646	3	142	1	329	2	
8		988-995	960	6	442	3	774	4	514	3	
9		996-999	929	6	822	5	365	2	451	3	

The table above is a compact representation of a 40-hour sample. Thus, the demand in the first hour was for six units; in the sixth hour for five units; in the twelfth hour (second column, second row), for four units, etc.

Next, we assign Monte Carlo numbers to the consecutive delays, $A-B$ and $B-E$.

DELAY (HR) $A-B$	ASSIGNMENT	DELAY (HR) $B-E$	ASSIGNMENT
7	000-049	13	000-039
8	050-499	14	040-169
9	500-999	15	170-379
		16	380-644
		17	645-879
		18	880-984
		19	985-999



--- Alternative routing

▨ Back orders

□ Transactions in the waiting line

Reorder point $= \bar{z}_{L.T.} + K\sigma_{L.T.} = 48 + (2)\sqrt{48} \approx 62$ units

$$r_0 = \sqrt{\frac{2\bar{z}C_r}{C_c}} = \sqrt{\frac{2(6000)2}{4(.12)}} \approx 225 \text{ units}$$

S.O.H. (r_0) = 76 units

$$\text{Delay } A-B \begin{cases} 7 \text{ hr } .050 \\ 8 \text{ hr } .450 \\ 9 \text{ hr } .500 \end{cases}$$

$$\text{Delay } B-C \begin{cases} 3 \text{ hr } .250 \\ 4 \text{ hr } .500 \\ 5 \text{ hr } .250 \end{cases}$$

$$\text{Delay } B-E \begin{cases} 13 \text{ hr } .040 \\ 14 \text{ hr } .130 \\ 15 \text{ hr } .210 \\ 16 \text{ hr } .265 \\ 17 \text{ hr } .235 \\ 18 \text{ hr } .105 \\ 19 \text{ hr } .015 \end{cases}$$

$$\text{Delay } C-D \begin{cases} 5 \text{ hr } .800 \\ 6 \text{ hr } .200 \end{cases}$$

$$\text{Delay } D-E \begin{cases} 5 \text{ hr } .200 \\ 6 \text{ hr } .200 \\ 7 \text{ hr } .300 \\ 8 \text{ hr } .300 \end{cases}$$

FIG. 9-12 Flow diagram for data and materials flow simulation.

Since we have generated the demand for each of 40 hours, we can determine the arrival times of the units at B and at E. We will trace through just a few of these demands in Table 9-1 as they step through the system links. In the

first hour there is a demand for six units. The six units are dispatched by the branch warehouse if there is sufficient S.O.H. The issue transaction is then forwarded from the branch warehouse to the data-processing center.

TABLE 9-1

HOUR	DEMAND AT A	DELAY A-B		TIME AT B	DELAY B-E		TIME AT E
		M.C. No.	HOURS		M.C. No.	HOURS	
1	6	127	8	9	521	16	25
2	0	623	9	11	180	15	26
3	2	155	8	11	180	15	26
4	3	710	9	13	604	16	29
5	3	427	8	13	604	16	29
6	5	594	9	15	792	17	32
7	1	125	8	15	792	17	32
8	6	002	7	15	792	17	32
9	6	593	9	18	834	17	35
10	6	592	9	19	052	14	33
11	6	006	7	18	834	17	35
12	4	118	8	20	968	18	38
13	1	591	9	22	386	16	38
14	4	812	9	23	349	15	38
15	6	614	9	24	088	14	38
16	4	177	8	24	088	14	38
17	4	456	8	25	977	18	43
18	3	501	9	27	750	17	44
19	3	997	9	28	318	15	43
20	5	169	8	28	318	15	43
21	1	107	8	29	424	16	45
22	4	702	9	31	836	17	48
23	3	607	9	32	594	16	48
24	3	496	8	32	594	16	48
25	0	823	9	34	140	14	48
26	2	474	8	34	140	14	48
27	1	992	9	36	126	14	50
28	1	209	8	36	126	14	50
29	4	353	8	37	331	15	52
30	2	833	9	39	899	18	57
31	1	129	8	39	899	18	57
32	1	002	7	39	899	18	57
33	2	872	9	42	960	18	60
34	2	768	9	43	285	15	58
35	3	942	9	44	050	14	58
36	4	454	8	44	050	14	58
37	2	353	8	45	377	15	60
38	2	315	8	46	130	14	60
39	3	499	8	47	560	16	63
40	3	935	9	49	035	13	62

There is a transmission delay from *A* to *B* of 8 hours. Internal delay in the data-processing center; i.e., from *B* to *E* is 16 hours. Therefore, the processing is completed at time $1 + 8 + 16 = 25$ hours. At hour 9 there is a demand for six units and again at hour 11, an additional six units are demanded. The issue transaction for these two, separate demands arrive at *B* at exactly the same hour—hour 18. Therefore, a single Monte Carlo number is used for hour 18 at *B* to determine the delay in transmission from *B* to *E*. This is then a characteristic which we are assuming for our data-processing

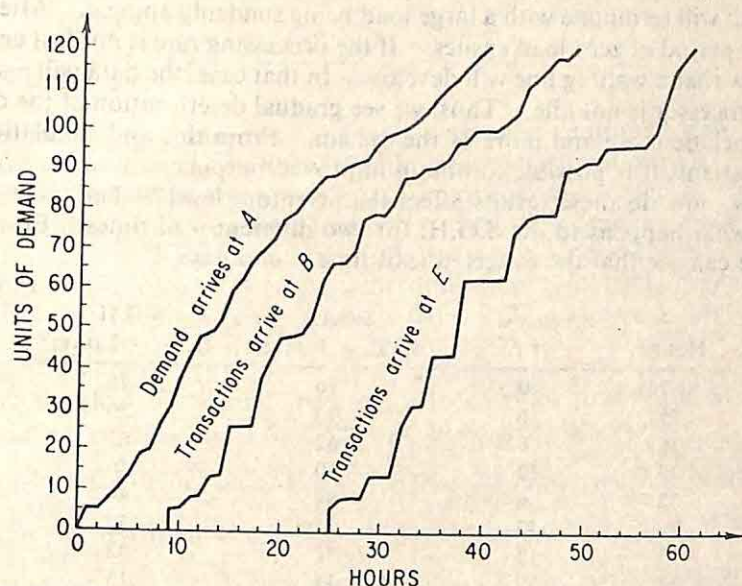


FIG. 9-13 Simulation of the cumulative load experienced by each processing department for the flow pictured in Fig. 9-12.

unit. Each hour has its own delay. Other possibilities might be tried. For example, each transaction could be assigned its own Monte Carlo numbers in which case the two arrivals at *B* might have different arrival times at *E*. Another variant would be to make delays a function of the load on the processing unit. Thus, if one transaction arrived at *B* at a particular hour, it could be quickly transmitted whereas, two, three, and more transactions would have progressively longer transmission times. With our simulation technique a large number of possible conditions can be explored. The assumption is made, of course, that controls exist which will permit each of the systems that are tried to be implemented. There is no advantage in simulating a system which is physically unrealizable. Without the ability to regulate the real system in accordance with the characteristics embodied in the simulation model, the matter becomes purely academic. We can now summarize our results, and they are of two kinds. First, the load characteristics as they are transmitted through the system. Second, the effect of these transmissions on the inventory levels.

Figure 9-13 presents the load characteristics as they arrive at *A*, *B*, and *E*.

We see that the curves are separated by varying amounts of time. These horizontal spans between points represent the expected delay plus and minus the particular delay variation which was observed. Thus, the total set of such horizontal lines between any two stations can be used to produce the simulation sample of the delay distribution. As we had previously predicted the curves become more jagged as we move from the earlier to the later steps in the network. Curve *E* has many plateaus each of which signify idle-time. Any plateau represents a period of time during which the load is zero. Then a plateau will terminate with a large load being suddenly applied. After this, another period of zero load ensues. If the processing rate is not fast enough, we know that a waiting line will develop. In that case, the data will age even if the processor is not idle. Thus, we see gradual deterioration of the output as we include more and more of the system. From this and simulations of other variants it is possible to obtain improved outputs.

Now, how do these results affect the inventory level? The table below shows what happens to the S.O.H. for two different lead times. From Fig. 9-12 we can see that the expected lead time is two days.

HOURS	ARRIVALS AT <i>E</i>	S.O.H. (L.T. = 4 DAYS)	S.O.H. (L.T. = 2 DAYS)
0-24	0	76	76
25	6	70	70
26	6	62	62
29	12	50	50
32	6	44	44
33	12	32	32
35	12	32	32
38	19	13	13
43	12	1	1
44	3	-2	-2
45	1	-3	222
48	12	-15	210
50	2	-17	208
52	4	-21	204
57	4	-25	200
58	9	-34	191
60	6	-40	185
61	0	185	185
62	3	182	182
63	3	179	179

Although the *Q*-model was designed to protect us against stock-outages in about 95 out of 100 lead time periods we observe that our simulation results would not bring joy to the inventory manager. On the contrary, he would ruefully observe that even when the lead time period would be exactly as expected, that a stock-outage results. With a lead time of four days (the probability of four days or more is about .05) the system is seriously malfunctioning. One order (indicated by the rectangles above) was required. (Stock-outs are shown as negative quantities and an oval indicates deliveries.)

What factor accounts for the poor showing? The answer is clear enough. The inventory rules were improperly designed. *They did not take the delay into account.* Whatever inventory model is used, it is necessary to incorporate delays required for data-processing and decision-making. Thus, the correct model would have been:

$$\text{Reorder point} = \bar{z}_{L.T. + \delta} + K(\sigma_{L.T. + \delta})$$

The quantity x_0 is unchanged, and reserve stock is larger by the amount

$$K\sigma_{L.T.} \left(\frac{\sqrt{L.T. + \delta} - \sqrt{L.T.}}{\sqrt{L.T.}} \right)$$

The total delay we remember increases with each additional step. Consequently, it should now be evident that the processing organization can play an important and determining role in the inventory cost system. This is one more example of why we have been stressing systemic costs. An additional factor exists. It is the variance of the total delay period itself. Just like lead time when it varies, we have a distribution of total delay periods. And just like lead time, when we choose a value $L.T._{\max}$, we may now choose a value, $\delta_{\max}$.

Let us calculate the new operating values for the model—and at the same time determine how much the delay is costing us. We have:

Expected delay (A-B) = 8.45 hr	Delay variance (A-B) = .345 hr
Expected delay (B-E) = 15.90 hr	Delay variance (B-E) = 1.870 hr
Expected delay (A-E) = 24.35 hr	Delay variance (A-E) = 2.215 hr

$$L.T._{\max} = 2 \text{ days} = 16 \text{ hours}$$

$$\delta_{\max} = 25 \text{ hours}$$

$$\text{Total} = 41 \text{ hours}$$

$$\text{Demand (in } L.T._{\max} + \delta_{\max}) = 123 \text{ units}$$

$$\sigma \text{ (in } L.T._{\max} + \delta_{\max}) = \sqrt{123} \approx 11 \text{ units}$$

$$\text{Reorder point} = 123 + (2)11 = 145 \text{ units}$$

The reserve stock has increased from 14 units (without transmission delays) to 22. This increase of eight units represents a cost of:

$$4(.12)(8) = \$3.84 \text{ per year}$$

A small change, but it does represent \$3840 to our 1000-item company. Further, it increases with price, and it can be much larger if the system compounds many delay variances. In that case, $\delta_{\max}$ must really be adjusted upward to take account of the long periods of delay. Furthermore, if

processing division *E* is actually idle when no load is applied—which means it has the capacity to handle and quickly dispatch the work load pulses as they arrive, then the costs of idle time of the processing divisions will be significant.

To conclude this section, let us see how our new reorder rule fares:

HOURS	ARRIVALS AT <i>E</i>	S.O.H. (L.T. = 2 DAYS)	S.O.H. (L.T. = 4 DAYS)
0-24	0	159	159
25	6	153	153
26	2	151	151
29	6	145	145
32	12	133	133
33	6	127	127
35	12	115	115
38	19	96	96
43	12	84	84
44	3	81	81
45	1	80	305
48	12	68	293
50	2	66	291
52	4	62	287
57	4	58	283
58	9	49	274
60	6	43	268
61	0	268	268
62	3	265	265
63	3	262	262

No shortages have occurred which presumably was our objective. The system is now protected from outages resulting from processing delays. It is certainly clear that data-processing design interacts with the performance of the inventory model and must be considered when selecting inventory rules. It is also clear that some portion of study funds should be allocated to design operating procedures. The inventory system as a practical entity cannot survive on a diet of pure inventory theory alone.

78. REGULATED INVENTORY SYSTEMS

A well-designed regulator will adjust and balance the flows between the components of a system. It will be totally *goal-oriented* in doing this. In other words, it is implicit that we must know what we are regulating and toward what end we are regulating. Ideally we would ask that the regulator minimize cost or maximize gain. But the problem of unifying all of the elements and setting down the rules whereby the probability of achieving the goals is maximized is usually inordinately difficult. A simple regulator cannot satisfactorily cope with the higher levels of system complexity which are inherent in the macroscopic view of an organization. Consequently, the

concept of total self-regulation has been evolved to meet the requirements imposed by complex systems with precisely defined goals. These more recent developments in control theory are based upon goal orientation and the difference between the state that exists and the state that is desired. The measured difference is the error. Decision rules are designed as functions of the error. The system's goal can therefore be stated as the optimization of some particular function of the error term. Error regulation is the core concept of missile systems, automatic pilots, and a variety of "true" automata. Given an error regulated system, further differentiation can be made as to the type of system and its level of sophistication. Definitions at this point become somewhat arbitrary, particularly when we start talking about servo-mechanisms and "true" automata.

Many systems possess characteristics which give them the appearance of being automatic while in fact they are only mechanized. Such systems appear to be automatic in the sense that a vending machine which issues an item when it is properly addressed appears to be automatic. Machines of this type can be very complex, dispensing many different kinds of items and accepting many different denominations of money for which they can make the correct change. However, our notions of automatic behavior have advanced far beyond this point and the vending machine is relegated to the ranks of mechanized equipment. The inventory models and systems which have been dealt with up to this point possess varying degrees of mechanization. These models have been designed to accept input signals (such as demand and lead time). In fact, they have been pre-designed with the expectation that certain kinds of signals will be fed into the system. (Compare this with the coin changer.) If the expectation is false; e.g., lead time is five times as great as expected, then the system lacks any capacity for adjusting itself to the new conditions. It is not designed to produce satisfactory outputs under any conditions that might occur. It lacks adaptive behavior.

The type of system which we are now introducing possesses adaptive behavior and fits our present-day conception of an automatic system. References have been made to it previously in the text but because of its predominant systems orientation its exposition was postponed until Part 2.

Because it is a meta-model with less practical significance and more conceptual importance, we did not wish to include this model with our other inventory models. Furthermore, it represents a synthesis of many elements—not the least of which is the notion of a regulator. Lastly, this system captures the essence of the relationship between the decision system and the operating system. Therefore, we feel that the greatest utility is derived by having postponed this discussion until the present time.

Many factors differentiate this inventory model from all of those which were previously discussed. We will mention a few of these differences. First, the basic servo-control model originated in the electrical engineering and electronics field. No other inventory model can claim this parentage. The primary application of the model remains in the originating field where it underlies the design of self-guiding automata and adaptive control systems. Second, the inventory adaptation of the control model is less important for its

operational characteristics than for its conceptual contributions.* Third, this is the only model discussed which *can be* totally self-regulating.

Refer to Fig. 8-3 and note that the executive and the inventory study are located within the control box. Our servo-model has an automatic control device located within that box. Therefore, our new system will be entirely self-contained and automatic. But the differences do not stop there. A fourth difference is the fact that the servo-system measures the discrepancy between the input and the output and adjusts its own rule box so as to produce a desired pattern of behavior. It does not compare the output to some internally stored criterion. Instead, it is dynamically responsive to a wide range of inputs and utilizes the feedback principle to produce the required modification of the outputs. Herein lies the major element of these control systems. Rules exist for monitoring the error signal and for transforming this measure of discrepancy. An efficient feedback channel is required to convey information about the output to the error sensor; thence to the control unit for transformation of the error signal and ultimately back into the system to modify the output. The modified output is then again fed back and compared with the input. A new error term is measured and is used to further modify the output. The process continues indefinitely—that is, until the control unit achieves its inherent objectives. The system then waits for any disturbance, such as a change in the input, to restimulate its control function.

The key to the system's performance obviously lies in the rules by which the control unit functions. So unlike our previous models which are static in spite of the fact that they have error-sensing and feedback, the servo-system is dynamic. The rules not only express the system's objectives but they incorporate that which has occurred with that which is desired. Figure 9-14 depicts a regulated inventory feedback control system. The symbols we will be using are explained in Fig. 9-14. It will be noticed that the input is shown as demand and the output can either be production or the quantity ordered. It is apparent that the difference between the input and the output represents the increase or decrease in inventory level which occurs at time t . Thus, the error term measures the rate of change of the inventory level. It then follows that the sum of the error terms over time is equal to the inventory shortage or the back-order level at time t . Let us take advantage of the clarity which can be obtained from the simulation of an example.

We will assume that there is constant demand for two units per day. On the third day of the week the demand unexpectedly doubles. Our control rule will be:

$$P_{t+1} = f(\epsilon_t) = \left(\frac{1}{2} \epsilon_t + P_t \right) = \frac{1}{2} P_t + \frac{1}{2} D_t$$

* The creator of this ingenious model is Herbert Simon. His original work, "On the Application of Servomechanism Theory in the Study of Production Control," appeared in *Econometrica*, 20, 2 (April, 1952), pp. 247-268. It still remains one of the most fundamental and well-organized presentations available in spite of the intervening years. Credit for further developments of significance in this area should be given to Ronald Howard and Robert Fromke. See R. Howard, *Notes on Operations Research*, 1959 (Cambridge, Mass.: The Technology Press, 1959), Chap. 5. Fromke's work, circa 1955, can be found in his unpublished dissertation, Columbia University.

That is, the control unit produces a signal* equal to one-half of the quantity ordered and received in stock during period t plus one-half of the units demanded during the interval. Our system is initially in equilibrium, which

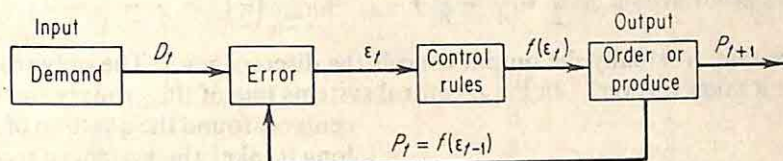


FIG. 9-14 Schematic of the regulated inventory feedback control system.

D_t = input signal at time t = demand in period t

P_t = output signal at time t = units produced in period t

ϵ_t = error measured at time t

$\epsilon_t = (D_t - P_t)$

$f(\epsilon_t)$ = modified error signal at time t

$P_{t+1} = f(\epsilon_t)$

implies that we produce or order two units per day. (The initial values of the demand and production can have any size including zero. Many times, for analytical convenience, the system is considered to begin at zero but this simplification is not required for our simulation.) The table which follows gives the steps of the simulation.

t	D_t	P_t	ϵ_t	$\frac{1}{2}\epsilon_t$	$f(\epsilon_t) = P_{t+1}$
1	2	2	0	0	2
2	2	2	0	0	2
3	4	2	2	1	3
4	2	3	-1	-.5	2.5
5	2	2.5	-.5	-.25	2.25
6	2	2.25	-.25	-.125	2.125
7	2	2.125	-.125	-.0625	2.0625

We could have continued for values of $t > 7$. But the pattern has been established and further iteration will produce:

- (1) continued reduction of the inventory shortage, and
- (2) gradual return of the production rate to the initial conditions.

Figure 9-15 charts the system we have just simulated. When the demand rose to four units, a back-order would result unless reserve stock is carried. The question we wish to ask is, can we be certain that the system will produce enough output stock to cover the pulse of demand which was experienced? The answer is yes, and we can demonstrate this as follows. Two units in excess of expected demand were required. (Units can of course represent hundreds, thousands, etc., of units.) The output generates a power series of

* A great variety of mechanical, electrical, and electro-mechanical devices are available to measure error and then transform it in accordance with the control rule, $f(\epsilon_t)$. For example, see W. G. Holzbock, *Automatic Control* (New York: Reinhold Publishing Corp., 1958).

additional stock produced. Thus, in period 3, one unit was produced above the normal level. In period 4, one-half unit; in period 5, one-fourth unit, etc. Thus,

$$\text{Output} = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{1 - \frac{1}{2}} = 2$$

We see that eventually the output cancels the discrepancy. The only trouble is that it takes forever. In these control systems one of the primary interests

centers around the question of how long it takes the system to recover—if it recovers. A better systems design—all other things being equal—is one in which recovery is rapid. But recovery may not result. If our system is satisfactory, then the sum of the error terms will be zero. It should be remembered that each error term is the inventory shortage or surplus which is created for that period of time. Since we are measuring the error as $D_t - P_t$ which is [demand (t) - production (t)], a positive error represents an inventory shortage while a negative error represents an inventory surplus. Then as we did for the output series,

Inventory shortage =

$$2 - \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 0$$

so our system at infinity has eliminated the inventory shortage.

Different input forms can occur. Our first example dealt with a demand pulse. The input increased for one period and then returned to its initial value. Many other

forms could be taken by the disturbance of the status quo. We could have a step increase. This would mean that the demand, initially at 2, would go to 4 and then stay at this higher level; a trend increase or decrease could occur. Thus, the system begins at 2 and in succeeding periods it experiences 3, 4, 5, 6, etc., or perhaps, 1.5, 1.25, 1.125, etc. The last input form we will mention, although there can be many others, is the seasonal type of change. Similar to a sinusoidal wave form, the sequence might be 2, 4, 6, 4, 2, 4, 6, etc.

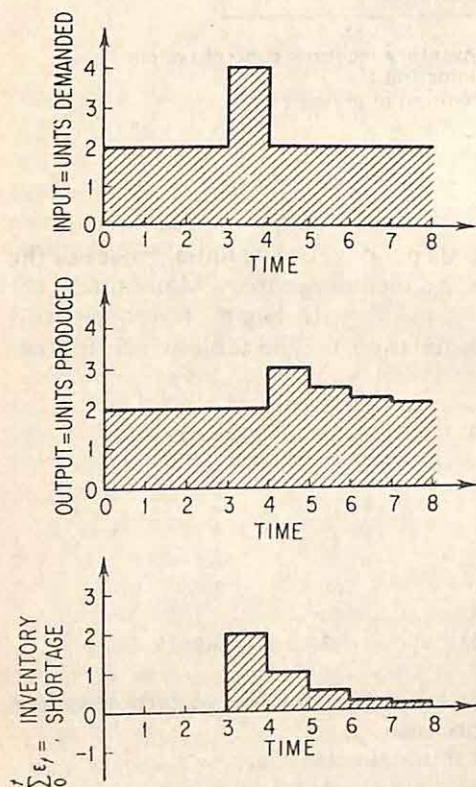


FIG. 9-15 Regulated inventory feedback system with control rule: $f(\epsilon_t) = (\frac{1}{2}\epsilon_t + P_t)$ and pulsed input.

Our system will respond differently to different input forms. Thus, the response and recovery of a system is now being considered in terms of basic input changes and not just random fluctuation. We will run a simulation, using our present decision rule but the form of the demand disturbance will be a trend of increasing demand. Let us see how our system fares under these conditions.

t	D_t	P_t	ε_t	$\frac{1}{2}\varepsilon_t$	$f(\varepsilon_t) = P_{t+1}$
1	2	2	0	0	2
2	2	2	0	0	2
3	4	2	2	1	3
4	6	3	3	1.5	4.5
5	8	4.5	3.5	1.75	6.25
6	10	6.25	3.75	1.875	8.125
7	12	8.125	3.875	1.9375	10.0625

Not very well! The output continues to diverge from the input. The sum of the errors increases and does not go to zero as we would like. Figure 9-16 represents the graphic portrayal of these results. If a trend (or ramp) input can occur, the system decision rule we have formulated is decidedly unsatisfactory.

We have now observed some major weaknesses in this particular form of decision rule. The fact that error-sensing and feedback were used did not prevent catastrophic results. Let us now translate this discussion into more specific inventory language. Consider the following model.

$$P_{t+1} = M - E[I] + \varepsilon_t = x_o + \varepsilon_t$$

where M is the maximum stock level as previously defined for the P -model; $E[I]$ is the expected inventory level, (S.O.H. + S.O.O.), at each reorder date; —so, $M - E[I]$ equals the expected order quantity, x_o . When this control model receives a pulse input, a continuous, undamped oscillation is set up in the system. Figure 9-17 illustrates the result derived in the table below.

t	D_t	P_t	ε_t	x_o	P_{t+1}
1	2	2	0	2	2
2	4	2	2	2	4
3	2	4	-2	2	0
4	2	0	2	2	4
5	2	4	-2	2	0
6	2	0	2	2	4
7	2	4	-2	2	0
8	2	0	2	2	4

From any point of view, this is not a happy result. We will therefore move on to a more sophisticated kind of regulator. It is still a regulator and not a self-regulator. (The latter is discussed in Section 79.)

$$P_{t+1} = M - E[I] + \sum_0^t \varepsilon_t = x_o + \sum_0^t \varepsilon_t$$

Examination of this equation will show that it embodies all of the characteristics of a P -model, where the reorder cycle is one unit of time and $E[I]$ incorporates both S.O.H. and S.O.O. The fact that control is now exercised as a function of an integrated error is clearly a major design difference.

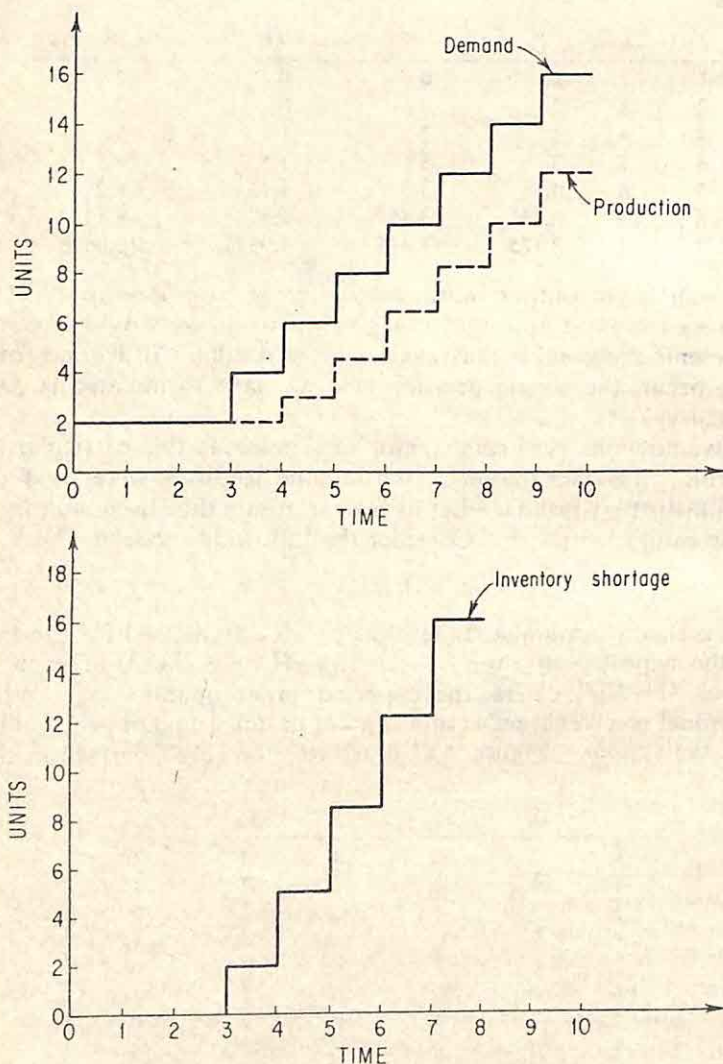


FIG. 9-16 Regulated inventory feedback system with control rule: $f(e_t) = [\frac{1}{2}e_t + P_i]$ and increasing input trend.

We already know how the P -model behaves when a pulse of demand occurs above the level of the expected demand. The reserve stock has been provided to protect against inventory shortages caused in this manner.

The reserve stock is brought back to its original level by increasing the next order that is placed. This was observed in our simulation of the P -model in the previous chapter. We will now want to investigate the behavior of the P -model when the input signal generated by our demand generator is no longer derived from a stationary distribution. In other words, what can we say about P -model performance when the mean of the demand distribution increases or decreases by some fixed amount? This is the input form which

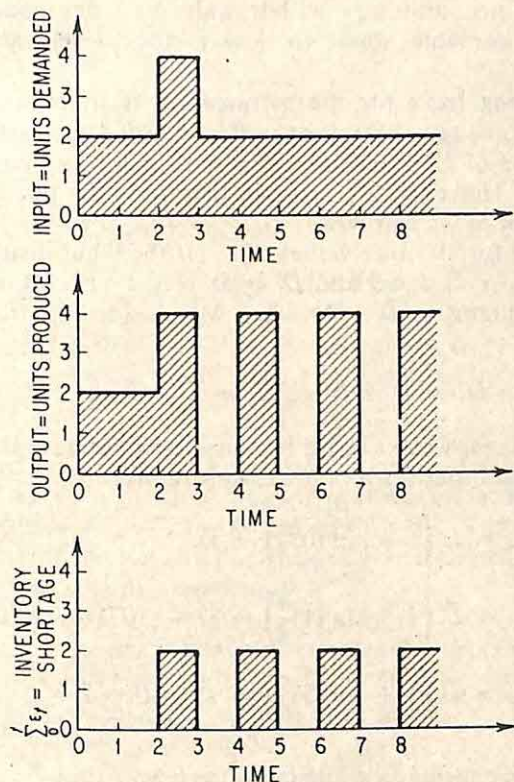


FIG. 9-17 Regulated inventory feedback system with control rule: $f(\epsilon_t) = x_0 + \epsilon_t$ and pulse input.

we called a step function. How does the P -model fare when the expected demand grows over time? We called this a trend input function. We know that our first model, not based on the summation error term, performed unsatisfactorily with the trend input form. The reader can ascertain for himself that it would have also been unsatisfactory with the step input form. But rather than run one simulation after another to find out how the P -model fares, we can use a mathematical method which transforms our difference equation and makes it quite simple to investigate the effects of different input forms on the behavior of the P -model. Our objective is to follow the inventory level over time. As in all systems analysis we must know what our

objectives are before we can get any benefits. The method we apply is known to control engineers as the z -transform. (Henceforth, for reasons explained in the next footnote, we will call this a z -transform.) It is specifically designed to resolve problems in a sample-data system. Our system is a sample-data system for the following reason. Only at discrete intervals is information about the stock level available. It is as though a relay closed once every so many time periods. When the relay is closed, information is transmitted about the stock level. When the relay is open there is no flow of information. Thus, although withdrawals may be made frequently no information is available about the exact stock level except at reported intervals.

The underlying basis for the z -transform is quite simple. When we write D_t we mean that D takes on different values at each time period, t . If the progression of D_t values is erratic, there will be no convenient form of representation. However, if the input distribution is postulated as a pulse occurring only once we can write: $D_t = \bar{D} + \Delta$, for a single, specific value of t and $D_t = \bar{D}$ for all other values of t . If the input disturbance is a step, we can write: $D_t = \bar{D}$, $t < \tau$ and $D_t = \bar{D} + \Delta$, $t \geq \tau$. It is clear that when $t \geq \tau$ the input jumps to $\bar{D} + \Delta$: when we consider the trend form for the input we have:

$$D_0 = \bar{D}, \quad D_1 = \bar{D} + \Delta \quad D_2 = \bar{D} + 2\Delta \dots D_m = \bar{D} + m\Delta$$

Lastly, although many cases could be considered, we have the seasonal variation or sinusoidal input form. Here, for example:

$$D_0 = \bar{D} \left(1 + \sin(0) \frac{\pi}{4} \right) = \bar{D}$$

$$D_1 = \bar{D} \left(1 + \sin(1) \frac{\pi}{4} \right) = \bar{D} + .707\bar{D} = 1.707\bar{D}$$

$$D_2 = \bar{D} \left(1 + \sin(2) \frac{\pi}{4} \right) = \bar{D} + \bar{D} = 2\bar{D}$$

and so forth.

Each input form follows a regular pattern which can be represented in the form of a power series. Thus:

Pulse:

$$F(t) = \Delta, 0, 0, 0, \dots$$

$$F(z) = (1)z^0 + (0)z^1 + (0)z^2 + \dots$$

Step:

$$F(t) = \Delta, \Delta, \Delta, \Delta, \dots$$

$$F(z) = (1)z^0 + (1)z^1 + (1)z^2 + \dots$$

This has the well-known convergence

$$F(z) = \frac{1}{1-z}$$

Trend:

$$F(t) = 0, \Delta, 2\Delta, 3\Delta, 4\Delta, \dots$$

$$F(z) = (0)z^0 + (1)z^1 + (2)z^2 + \dots$$

$$= z[(0)z^{-1} + (1)z^0 + (2)z^1 + (3)z^2 + \dots]$$

It will be noted that the series within the brackets on the right-hand side is the derivative of the step-input series. Therefore, we take the derivative

$$\frac{d}{dz} \left[\frac{1}{1-z} \right] = \left[\frac{1}{(1-z)^2} \right]$$

and

$$F(z) = z \left[\frac{1}{(1-z)^2} \right] = \left[\frac{z}{(1-z)^2} \right]$$

Exponential:

$$F(t) = 1, e^{-a}, e^{-2a}, e^{-3a}, \dots$$

$$F(z) = (1)z^0 + (e^{-a})z^1 + (e^{-2a})z^2 + \dots$$

We can write the closed form directly:

$$F(z) = \frac{1}{1 - e^{-a}z} = \frac{e^a}{e^a - z}$$

By using the method above it is possible to obtain z -transforms for all of the situations that we will encounter. An abbreviated table of z -transforms is given below. We will use it directly for all of the developments which follow. One of the advantages of these transforms lies in their easy transferability from a table to an equation.

SHORT TABLE OF z -TRANSFORMS*

ORIGINAL FUNCTION	z -TRANSFORM
P_t (function of t)	$P(z)$
P_{t+1} (one period delay)	$P(z)/z$
P_{t-k} (k periods of delay)	$P(z)/z^k$
$\sum_0^t P_t$ (summation)	$P(z)/(1-z)$
Unit impulse, $t = 0$	1
Unit step, a constant	$1/(1-z)$
Trend, increases with t	$z/(1-z)^2$

* We have been using the generating function

$$\sum_{n=0}^{\infty} y_n z^n = F(z)$$

and calling it the z -transform. The z -transform would be obtained by

$$\sum_{n=0}^{\infty} y_n z^{-n} = F(z)$$

The resulting z -transforms, familiar to electrical engineers, are slightly different from those shown in the table.

We will now transform the equation for the P -model regulator; the output will be considered as being measured from the x_0 level as a base.

$$P_{t+1} = S_t = \sum_0^t \varepsilon_t = \sum_0^t D_t - \sum_0^t P_t$$

Then

$$\frac{P(z)}{z} = S(z)$$

where S_t = the inventory shortage at t .

And

$$\frac{P(z)}{z} = \frac{D(z)}{(1-z)} - \frac{P(z)}{(1-z)}$$

With respect to the last equation, we can simplify. Thus:

$$P(z) = zD(z) \quad \text{and} \quad S(z) = D(z)$$

These are our basic equations. We will now apply several different input forms in order to observe the model's behavior.

(1) *Impulse*: Demand distribution remains unchanged and pulse represents a deviation from the expected value.

$$D(z) = 1 \quad P(z) = z \quad S(z) = 1$$

We have $P(z)/z = 1$. Therefore, $P_1 = 1$. The impulse occurs when $t = 0$; the output increases just enough to compensate in the following period. Also: $S(z) = 1$. Therefore, $S_0 = 1$. At the time of the demand an inventory shortage occurs. For all subsequent periods the inventory shortage returns to zero. This is precisely the behavior we want and expect from the P -model. The pulse bites into reserve stock which is replenished in the following period. (See Fig. 9-18.)

(2) *Step*: The demand distribution shifts to a new expected value at time t .

$$D(z) = \frac{1}{1-z} \quad P(z) = \frac{z}{1-z} \quad S(z) = \frac{1}{1-z}$$

We have $P(z)/z = 1/(1-z)$ which is the unit step. Therefore, $P_{t+1} = 1$. This means that one period after the demand change occurs the output rises to one and stays there. Also, $S_t = 1$. The inventory shortage which occurs when the demand is made is *never* rectified. (See Fig. 9-18.) This is obviously not a situation to be desired. It clearly points out the need for updating the inventory model parameters in some regular fashion. We must bear in mind that although we are talking about an inventory shortage, this shortage is applied to the reserve stock level. Thus, when a shift in demand occurs it will not be as easily recognized as if a real shortage developed. The effect may be hidden, expressed only in a higher percentage of stock-outages. Thus, the cost of updating is offset by the increased cost of outages.

For highly stable demand generators, the need for frequent review is obviously less important. This problem becomes even more compelling for the next two input forms.

(3) *Trend*: The demand distribution shifts to increasingly higher or lower levels.

$$D(z) = \frac{z}{(1-z)^2} \quad P(z) = \frac{z^2}{(1-z)^2} \quad S(z) = \frac{z}{(1-z)^2}$$

We have $P(z)/z = z/(1-z)^2$ which is the linear trend. Thus, one period

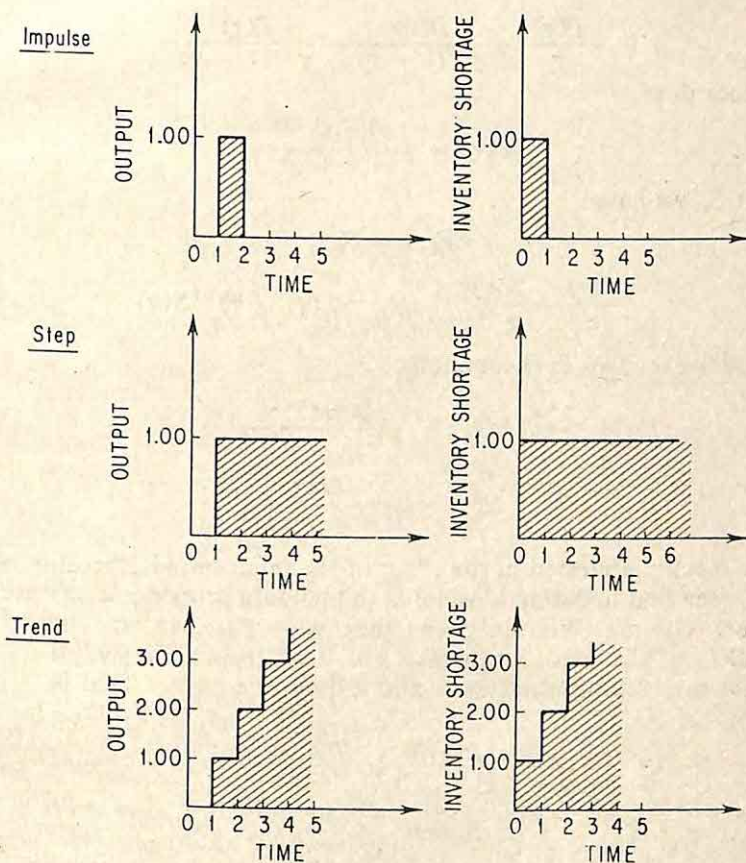


FIG. 9-18 P-model regulator.

after demand goes to one, the production goes to one. At the second period it goes to two, etc., that is, $P_0 = 0$, $P_1 = 1$, $P_2 = 2$, $P_3 = 3$, $\dots$, $P_t = t$. Meanwhile, the inventory shortage is also following the trend line, beginning

at $t = 0$. The reserve stock will be completely eaten up, plunging the system into back-orders and lost customers. (See Fig. 9-18.)

(4) *Delay*: We have been patiently waiting to discuss the problems occasioned by delay but before we could get to it the basic material had to be developed. We will assume that the demand distribution is stationary. There is no step or trend shifting. However, a demand which takes place at time t is delayed in processing. The period of delay is k . Thus, our model now becomes:

$$P_{t+1} = \sum_0^{t-k} D_{t-k} - \sum_0^{t-k} P_{t-k}$$

The transforms will be found in the table we have given.

$$\frac{P(z)}{z} = \frac{D(z)}{z^{-k}(1-z)} - \frac{P(z)}{z^{-k}(1-z)}$$

We reduce this:

$$P(z) = \frac{D(z)z^{k+1}}{1-z+z^{k+1}}$$

And for S_t , we have:

$$P_{t+1} = S_{t-k}$$

Thus:
$$\frac{P(z)}{z} = \frac{S(z)}{z^{-k}} \quad \text{or} \quad P(z) = z^{k+1}S(z)$$

We combine the two $P(z)$ equations:

$$z^{k+1}S(z) = \frac{D(z)z^{k+1}}{1-z+z^{k+1}}$$

or

$$S(z) = \frac{D(z)}{1-z+z^{k+1}}$$

We are mainly interested in the effect of the input impulse because we have already seen that updating is essential to maintain proper cost balance in the inventory system. What happens then with different amounts of delay when $D(z) = 1$? First, let us take $k = 0$ which is no delay. This should yield our original impulse system and it does of course. That is,

$$P(z) = \frac{zD(z)}{1-z+z}$$

and

$$S(z) = \frac{D(z)}{1-z+z}$$

Now, let us set k equal to one. We obtain the two equations:

$$P(z) = \frac{z^2}{1-z+z^2}$$

and

$$S(z) = \frac{1}{1-z+z^2}$$

These transforms are not in our table, but they can be quickly obtained by using long division to obtain a power series thus:

$$\begin{array}{r}
 1 - z + z^2 \overline{) 1 + z - z^3 - z^4 + z^6 + \dots} \\
 \underline{1 + 0 + 0 + 0 + 0 + 0 + 0 + 0} \\
 1 - z + z^2 \\
 \underline{+ z - z^2 + 0} \\
 z - z^2 + z^3 \\
 \underline{- z^3 + 0 + 0} \\
 - z^3 + z^4 - z^5 \\
 \underline{- z^4 + z^5 + 0} \\
 - z^4 + z^5 - z^6 \\
 \underline{+ z^6 + 0 + 0} \\
 z^6 - z^7 + z^8
 \end{array}$$

Let us write out this power series and then interpret it:

$$S(z) = (1)z^0 + (1)z^1 + (0)z^2 + (-1)z^3 + (-1)z^4 + (0)z^5 + (1)z^6 \dots$$

What it tells us is that the inventory shortage oscillates. At $t = 0$, it is one. At $t = 1$, it is one. At $t = 2$, it is zero. At $t = 3$, it is minus one. Meanwhile, the output series is the same as the above multiplied by z^2 .

$$P(z) = (0)z^0 + (0)z^1 + (1)z^2 + (1)z^3 + (0)z^4 + (-1)z^5 + (-1)z^6 + \dots$$

For convenience, we tabulate these results.

t	P_t	S_t
0	0	1
1	0	1
2	1	0
3	1	-1
4	0	-1
5	-1	0
6	-1	1
7	0	1
8	1	0

Figure 9-19 illustrates the endless oscillation which the system experiences.

What is the cost of this behavior? It can be very high indeed. Errors do not cancel out as far as inventory costs are concerned, and so we have outage costs as well as excess carrying costs to contend with. This is then a very great weakness of the P -model. It is extremely vulnerable to transmission and processing delays. That is why we will now go on to discuss a servo-control model which we shall call a super- P -model.

79. SELF-REGULATED INVENTORY SYSTEMS

The two weaknesses of the P -model are its lack of responsiveness to shifts and new trends in demand and second, its inability to suffer delays without

producing large oscillations. Of course, once these facts are recognized the *P*-model can be an efficient inventory model. Delays must be controlled and whatever delays exist must be designed so as to avoid causing oscillation.

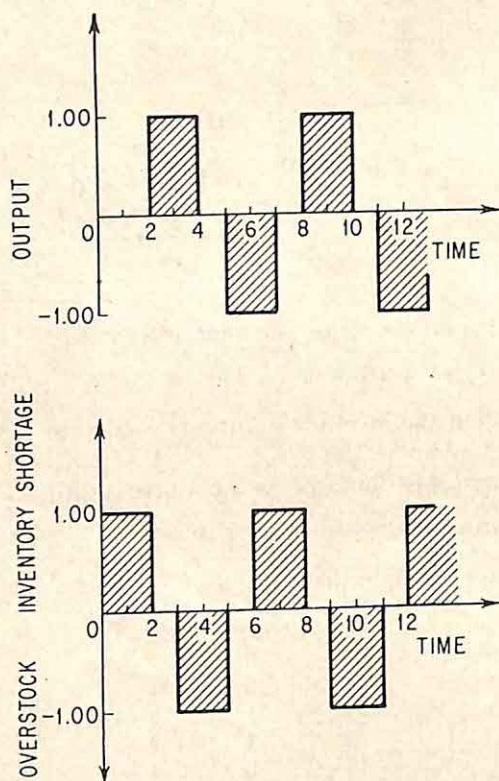


FIG. 9-19 *P*-model regulator with impulse input and one additional period for processing delay.

The reason that the *P*-model can't do a better job is the fact that M , the maximum level is invariant. It is in fact totally unresponsive to whatever is happening. That is why the *P*-model has been called static. However, the self-regulating principle eliminates this condition. Compare the two equations below:

$$P_{t+1} = M - E[I] + \sum_0^t \varepsilon_t \quad (1)$$

$$P_{t+1} = M - E[I] + \sum_0^t \varepsilon_t + \sum_0^t \sum_0^t \varepsilon_t$$

The first is, of course, our *P*-model. The second is the model we have named super-*P*. It will be seen that the second model has a double summation

term. It is the sum of the inventory shortage over time. Let us consider the dimensions of this measure. We have

$$\varepsilon_t = \text{units per day short}$$

$$\sum_t \varepsilon_t = \text{units short}$$

then

$$\sum_t \sum_t \varepsilon_t = \text{unit-days short}$$

So, this additional factor introduces a means of assessing outage penalties—not just for units short, but for units short for how long. (It will be remembered that we discussed such an outage penalty at a previous point.) Similarly, we not only have units in excess of the chosen level but we have excess units carried for some period of time. It is this time-unit measure which acts to modify the reference level, M . In fact, in the second model, the M level is M' , where

$$M' = M + \sum_0^t \sum_0^t \varepsilon_t$$

Thus, when the number of shortage days increases, the M' level rises with it. When the number of excess carrying days goes up, the M' level falls. What this really means, of course, is that the reserve stock has been made variable and a function of the time-units summation.

Now, we will make one additional change. We rewrite our self-regulator equation:

$$P_{t+1} = x_0 + A \sum_0^t \varepsilon_t + B \sum_0^t \sum_0^t \varepsilon_t$$

This gives us an additional element of control with respect to the self-regulating characteristic of M . The M -level can be interpreted as follows:

$$M' = M + A \sum_0^t \varepsilon_t + B \sum_0^t \sum_0^t \varepsilon_t - \sum_0^t \varepsilon_t = M + (A - 1) \sum_0^t \varepsilon_t + B \sum_0^t \sum_0^t \varepsilon_t$$

It will be noted that by choosing $A = 1$, $B = 0$, we get $M' = M$ which is the regulator P -model. Therefore, as would be expected the regulator P -model is just a special case of the self-regulating super- P -model.

Of particular importance is the case where $A = B = 1$. Let us obtain the $\mathcal{Z}$ -transform of the self-regulating model* in order to explain this case:

$$\begin{aligned} P_{t+1} &= A \sum_0^t \varepsilon_t + B \sum_0^t \sum_0^t \varepsilon_t \\ &= A \sum_0^t D_t + B \sum_0^t \sum_0^t D_t - A \sum_0^t P_t - B \sum_0^t \sum_0^t P_t \end{aligned}$$

Then:

$$\frac{P(\mathcal{Z})}{\mathcal{Z}} = \frac{AD(\mathcal{Z})}{1-\mathcal{Z}} + \frac{BD(\mathcal{Z})}{(1-\mathcal{Z})^2} - \frac{AP(\mathcal{Z})}{1-\mathcal{Z}} - \frac{BP(\mathcal{Z})}{(1-\mathcal{Z})^2}$$

* As before, the output is measured from the x_0 level as a base.

and
$$P(z) \left[\frac{1}{z} + \frac{A}{1-z} + \frac{B}{(1-z)^2} \right] = D(z) \left[\frac{A}{1-z} + \frac{B}{(1-z)^2} \right]$$

$$P(z) \left[\frac{(1-z)^2 + Az(1-z) + Bz}{z(1-z)^2} \right] = D(z) \left[\frac{A(1-z) + B}{(1-z)^2} \right]$$

Whence:

$$\begin{aligned} P(z) &= D(z) \left\{ \frac{z[A(1-z) + B]}{(1-z)^2 + Az(1-z) + Bz} \right\} \\ &= D(z) \left[\frac{(A+B)z - Az^2}{(A+B-2)z + (1-A)z^2 + 1} \right] \end{aligned}$$

When the coefficients are chosen equal to one the transform is reduced as follows:

$$P(z) = D(z)(2z - z^2)$$

and it will be found that:

$$S(z) = D(z)(1-z)$$

The respective power series with impulse input are:

$$P_t = (0)z^0 + (2)z^1 + (-1)z^2 + (0)z^3 + (0)z^4 + \dots$$

and
$$S_t = (1)z^0 + (-1)z^1 + (0)z^2 + (0)z^3 + (0)z^4 + \dots$$

So the inventory shortage becomes one unit immediately. In the period which follows an excess of one unit will be carried and thereafter the inventory returns to the desired level. Output jumps to two units in excess of normal output in the period following the withdrawal. It is then reduced to below normal output—by one unit—in the following period. Thereafter it returns to normal. This pattern of response is known as critical damping. It is the fastest response that the system can make without overdamping. On the face of it, one would usually prefer critical damping, but an analysis of the costs involved is required to be certain of this. Let us simulate a number of the different responses which this system will provide when A and B take on different values.

Case 1: *Divergent Oscillation* (See Fig. 9-20)

Input = Impulse

$$A = 2, \quad B = 1$$

$$P_{t+1} = x_0 + 2 \sum_0^t \varepsilon_t + \sum_0^t \sum_0^t \varepsilon_t$$

t	D_t	P_t	ε_t	$\sum_0^t \varepsilon_t$	$\sum_0^t \sum_0^t \varepsilon_t$	$2 \sum_0^t \varepsilon_t$	x_0	P_{t+1}
1	2	2	0	0	0	0	2	2
2	4	2	2	2	2	4	2	8
3	2	8	-6	-4	-2	-8	2	-8*
4	2	-8	10	6	4	12	2	18
5	2	18	-16	-10	-6	-20	2	-24*
6	2	-24	26	16	10	32	2	44
7	2	44	-42	-26	-16	-52	2	-66*
8	2	-66	68	42	26	84	2	112

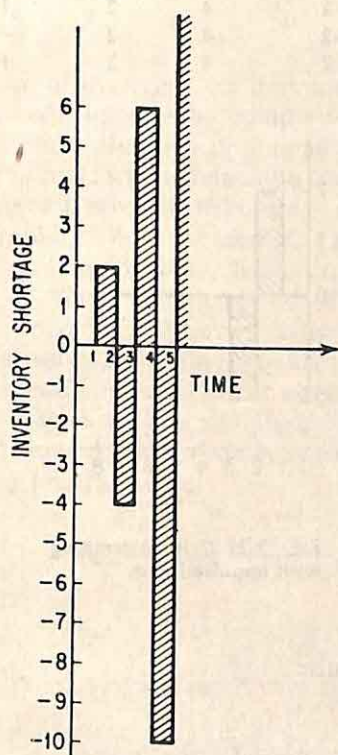


FIG. 9-20 Divergent oscillation with impulse input.

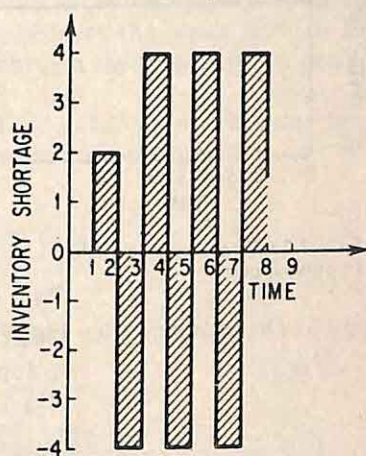


FIG. 9-21 Undamped oscillation with impulse input.

* To order a negative amount would be equivalent to dumping some of the stock on hand. For this to be feasible—let alone reasonable—there must be enough stock on hand. To obtain a reasonable simulation we would substitute zero for negative outputs. However, it is our desire to show the underlying response pattern so we have not used zero.

Case 2: Undamped Oscillation (See Fig. 9-21.)

Input = Impulse

$$A = 1, \quad B = 2$$

$$P_{t+1} = x_0 + \sum_0^t \varepsilon_t + 2 \sum_0^t \sum_0^t \varepsilon_t$$

t	D_t	P_t	ε_t	$\sum_0^t \varepsilon_t$	$\sum_0^t \sum_0^t \varepsilon_t$	$2 \sum_0^t \sum_0^t \varepsilon_t$	x_0	P_{t+1}
1	2	2	0	0	0	0	2	2
2	4	2	2	2	2	4	2	8
3	2	8	-6	-4	-2	-4	2	-6
4	2	-6	8	4	2	4	2	10
5	2	10	-8	-4	-2	-4	2	-6
6	2	-6	8	4	2	4	2	10
7	2	10	-8	-4	-2	-4	2	-6
8	2	-6	8	4	2	4	2	10

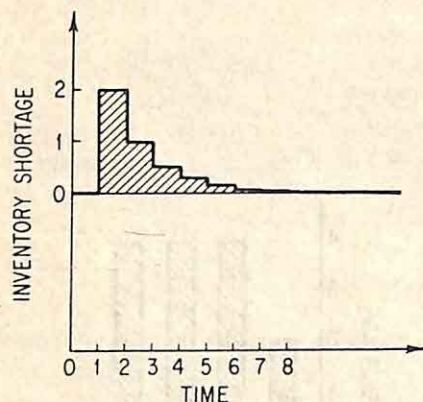


FIG. 9-22 Overdamping with impulse input.

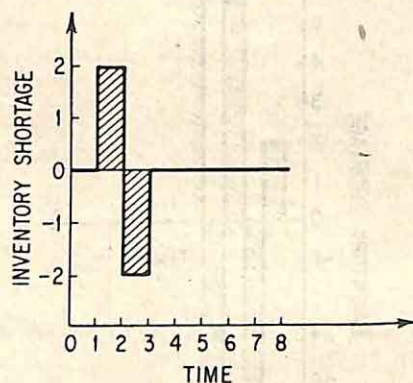


FIG. 9-23 Critical damping with impulse input.

Case 3: Overdamping (See Fig. 9-22.)

Input = Impulse

$$A = \frac{1}{2}, \quad B = 0$$

$$P_{t+1} = x_0 + \frac{1}{2} \sum_0^t \varepsilon_t + (0) \sum_0^t \sum_0^t \varepsilon_t$$

t	D_t	P_t	ε_t	$\sum_0^t \varepsilon_t$	$\frac{1}{2} \sum_0^t \varepsilon_t$	x_0	P_{t+1}
1	2	2	0	0	0	2	2
2	4	2	2	2	1	2	3
3	2	3	-1	1	$\frac{1}{2}$	2	$2\frac{1}{2}$
4	2	$2\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	2	$2\frac{1}{4}$
5	2	$2\frac{1}{4}$	$-\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{8}$	2	$2\frac{1}{8}$
6	2	$2\frac{1}{8}$	$-\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	2	$2\frac{1}{16}$

Case 4: Critical damping (See Fig. 9-23.)

Input = Impulse

$$A = 1, \quad B = 1$$

$$P_{t+1} = x_0 + \sum_0^t \varepsilon_t + \sum_0^t \sum_0^t \varepsilon_t$$

t	D_t	P_t	ε_t	$\sum_0^t \varepsilon_t$	$\sum_0^t \sum_0^t \varepsilon_t$	x_0	P_{t+1}
1	2	2	0	0	0	2	2
2	4	2	2	2	2	2	6
3	2	6	-4	-2	0	2	0
4	2	0	2	0	0	2	2
5	2	2	0	0	0	2	2

The case of divergent oscillation is most certainly an inventory manager's nightmare. Perhaps by comparison the undamped oscillation seems to be not so bad. But this is pure illusion unless the amplitude is very small. The amplitude will be related to the variance of the demand distribution or of the forecast error distribution. If this is large, then oscillations can be catastrophic. Both overdamping and critical damping seem to be permissible. The overdamping case, however, never quite makes up for the impulse demand. Over a long period of time the inventory shortage is reduced almost to zero. However, critical damping restores the stock level to its original value and rather quickly. That is why, on the face of it, we prefer critical damping to the other alternatives.

Now let us see how well the coefficients for critical damping serve us when an additional period of delay occurs. The general expression for k periods of delay i.e., P_{t+k} , is.

$$P(z) = D(z)z^k \left[\frac{(A+B) - Az}{1 + (A+B)z^k - Az^{k+1} - 2z + z^2} \right]$$

and
$$S(z) = D(z) \left[\frac{1-z}{1 + (A+B)z^k - Az^{k+1} - 2z + z^2} \right]$$

When $k = 2$ and the input is of the impulse type, we find:

$$P(z) = D(z)z^2 \left[\frac{(A+B) - Az}{1 + (A+B)z^2 - Az^3 - 2z + z^2} \right]$$

and
$$S(z) = D(z) \left[\frac{1-z}{1 + (A+B)z^2 - Az^3 - 2z + z^2} \right]$$

We can see that $A = B = 1$ will not reduce the denominators of these expressions and consequently a variety of oscillatory behaviors can occur. There is, in fact, nothing that we can do that is reasonable to prevent some loss from occurring.

This is why we have emphasized the importance of control over delays.

Without a systems study which includes data-processing and data-transmission as an integral part of the investigation, the anticipated behavior of the "improved" inventory system may surprise some and cause consternation for others.

80. FORECASTING AND UPDATING CONTROLS

Even a super-model will not provide the ideal kind of inventory system. The ordinary P - and Q -models are more vulnerable than the self-regulating type and so delays in the system must be carefully controlled and their existence must be justified by a balancing of costs. But how can we possibly protect the system against the variety of input forms which can arise?

The answer is that adequate forecasting and updating procedures go a long way toward eliminating the weaknesses and liabilities of these inventory models. Let us consider forecasting first.

The problem may be stated this way. Given an unknown demand generator, what can we find out about it? Usually we cannot open it up and study the way it works. If we have a contractual demand, we know exactly how the generator works. If the problem is one of maintaining the right number of spare parts for a machine, the demand generator would be known when a policy of preventive maintenance is followed, but less so if remedial maintenance is practiced. When the demand originates in the market place very little can be said about the construction of the demand generator that is really useful. How do we manage to forecast then? Generally, we bypass the structural problem. Instead, we attempt to observe the relationship between input and output. The brand name is an input. Advertising and promotional dollars, competitive behaviors, the state of the economy are inputs too. Figure 9-24 presents a schematic picture of the demand generator illustrating a few of the inputs to which it is usually responsive.

No matter how long we study a system only a small part of the total number of inputs can be isolated and studied. Fortunately, most systems are relatively unresponsive to the majority of inputs. This reduces the size of the problem which must be studied, but even if only a few inputs are found to play a major role, the problem is complicated by factors such as the existence of thresholds which when surpassed lead to extremal responses. Flat response systems can interact when certain combinations of the inputs occur. The system may be dynamic moving from one state of stability into another and never returning to its previous condition. The transitions could occur, for example, when one of the inputs takes on an extreme value, thereby permanently altering the demand system.

As we have noted, some demand systems do not appear to be particularly sensitive to inputs over most of the range of values which these inputs can take. They seem to be almost independent of the great variety of inputs to which they are exposed. When we consider that feedback networks are operating in the system, this fact does not seem at all surprising. A single pulse of promotional input can spark repurchase behavior and, in addition,

the consumer can become a source of self-generating input. He will bring friends and acquaintances into the demand generator long after the original, single input pulse. The existence of many such internal generators, which continue to operate with or without further external stimulation can result in a stable output that is relatively insensitive to external inputs. Systems of this type possess all of the characteristics of an inertial system. On the other

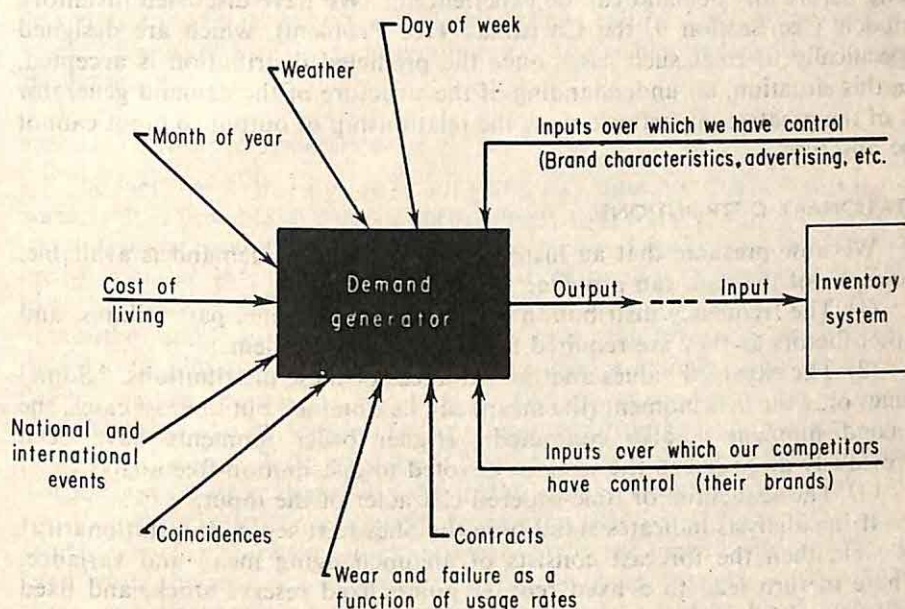


FIG. 9-24 The demand generator system.

hand, some systems are quite responsive to small changes in the inputs, but if enough inputs are present the effects may tend to cancel each other and the system can be considered to be stable. We will categorize this general class of distributions as being of a stable type. We will seldom know how the stability is achieved but we must know if it exists. And we must also know when it stops existing. For this purpose, the Shewhart criterion* as it is applied in statistical quality control practice can serve as an adequate definition of stability. Since this is a well-known criterion, available from many sources, yet far too subtle to be properly treated in a summary fashion we will not develop it here. At the same time, as we will demonstrate, the definition is arbitrary since other forms of stability exist. The problem is a semantic one, but the knowledge of this fact is not particularly helpful. In some ultimate sense we might define stability as "possessing some pattern of regular behavior." The Shewhart criterion is then one kind of stability.

* Walter A. Shewhart, *Statistical Method from the Viewpoint of Quality Control* (Washington, D.C.: The Department of Agriculture, 1939).

NO HISTORICAL INFORMATION IS AVAILABLE

If a sufficient history of past demand is not available then demand forecasts must be made on the basis of market potential studies, general surveys, and whatever parallel experiences can be obtained. These factors are then mixed together with select experiences, insights, and intuitions. The brew which results has been known to be quite accurate. Unfortunately, as many negative results can be reported. Essentially this is a special problem. It is of great importance in the style goods industry where orders must be placed long before any demand can be experienced. We have discussed inventory models (see Section 9, the Christmas Tree Problem), which are designed specifically to treat such cases once the predicted distribution is accepted. In this situation, an understanding of the structure of the demand generator is of the greatest importance since the relationship of output to input cannot be observed.

STATIONARY DISTRIBUTIONS

We now presume that an historical record of past demand is available. Analysis of the data can provide:

(1) The frequency distribution for demand, lead time, part failures, and other factors as they are required for a particular problem.

(2) The expected values and the variances of these distributions. Sometimes only the first moment (the mean) can be obtained but in most cases, the second moment is also computed. Higher order moments have been previously discussed in the sections devoted to distribution-free analysis.

(3) The sequential or time-ordered character of the input.

If the analysis indicates stability in the Shewhart sense, and stationarity* as well, then the forecast consists of an unchanging mean and variance. These in turn lead to a fixed reorder point, fixed reserve stock, and fixed M -level. In the sections devoted to regulators it became amply clear that changes of the step and trend types could not be adequately handled by our ordinary inventory models. These decision rules do not fulfill the objectives of the inventory study under any conditions except stationarity. If, on the other hand, the changes in the input can be anticipated, then dynamic inventory models can be used which optimize over some given planning horizon.

If the demand distribution which is generated is stationary, then it is possible to provide "good" forecasts. Naturally, that does not mean error-free forecasts. On the contrary, the distribution of errors will closely approximate the demand distribution. Thus, let the expected demand be $\bar{z}$ and the observed variance be σ_z . If the control chart indicates stability of these parameters—then the forecast, $F(z) = \bar{z}$. Also, if records are kept of the actual demand minus the expected demand, then the distribution of these

* By stationarity we mean that all of the parameters and the distribution itself are time-invariant. For a detailed discussion, see William Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1 (New York: John Wiley and Sons, Inc., 1950), pp. 328–329.

differences should have zero mean, $\bar{\varepsilon} = 0$, and $\sigma_{\varepsilon} = \sigma_z$. Further, if the mean demand is improperly estimated, then the error mean can be used to correct this forecast. But no such change should be made until the stability criterion is violated. This approach does not require the conscious utilization of the error distribution. Analysis of the data with control monitoring is sufficient to detect a basic change in the form of the input. Pulses are the normal expectation and in general they will be accepted by the control monitor. If a step or trend input occurs either in the mean or the variance it will be observed at an early stage and suitable alterations can be made in the decision rules. If such changes were to go unnoticed the system-wide penalties could be significant.

NONSTATIONARY DISTRIBUTIONS

The fact that a distribution is not stationary does not mean that it is not stable. It is possible to derive control charts with trend lines. In this case, the shift in the mean follows a consistent pattern and the anticipated demands are predictable in the full sense of statistical quality control. Allowance must be made in the design of the inventory decision rules to accept such variants of stable distributions and to optimize the inventory performance over the planning horizon. As the mean or variance changes, in an anticipated fashion, the reorder points, the M -levels and the reserve stock requirements should be altered. Deviations are now measured from the new mean or trend line. The reserve stock must be modified as the variance of the demand distribution changes along with the mean. If the shape of the distribution and therefore the variance is unchanging, then the reserve stock remains the same. There is no reason, for that matter, why the mean could not remain fixed while the variance experiences a step or trend change. In this latter case, since the reserve stock is modified—the reorder point or the M -level will be affected as well. A nonstationary but stable distribution requires, if anything, more attention than the stationary case. The fact that change is taking place increases the probability that a threshold will be passed or that an interaction at some level will occur.

NONSTABLE DISTRIBUTIONS

The real challenge in forecasting arises when an apparent lack of stability in the Shewhart sense is experienced. As we have noted previously in this section, stability can exist but in a form which does not lend itself to control chart procedures. For this case, a number of variants exist.

(1) The historical forecast presupposes a time of year stability. Thus, what happened in month j , a year ago, is considered the best estimate of what will happen in month $j + 12$.

(2) The historical forecast is a special case of the base series modification of an economic time series. Namely, the mean is considered to be unchanged. If, however, a shift in demand (step input) is detected, then the base series is applied to this new estimate. Thus, the base series of 59, 36, 40 would be $59/45 = 1.31$, $36/45 = .80$, and $40/45 = .88$, where the mean of the demands

is 45. Then, if a step is anticipated to a mean level of 50, the expected series would be $(50 \times 1.31 = 66.5)$, $(50 \times .80 = 40)$, and $(50 \times .88 = 44)$.

(3) A variety of techniques exist* for the general modification of a time series. In particular, it is frequently desired to remove the contributions of older values and give greater credence to the more recent events. For example, a four month moving average is calculated:

$$\frac{z_1 + z_2 + z_3 + z_4}{4}$$

and is continuously updated as follows:

$$\frac{z_2 + z_3 + z_4 + z_5}{4}$$

Another approach which is particularly well suited for computer applications is the use of exponential smoothing. With this method, only the old expected value is retained. We then have:

$$\bar{z}_{j+1} = \alpha z_j + (1 - \alpha)\bar{z}_j$$

where $\bar{z}_j$ is the expected demand in the j th period, z_j is the actual demand that was experienced in the j th period and α is the smoothing constant. Usually this constant is kept very small (about .10) in order to decrease responsiveness to random fluctuations.

(4) Correlational techniques are utilized when the demand series appears to be related to another time series which it lags. Thus, if a particular economic index (such as housing starts or cost of living, etc.) appears to lead the demand for our products, a lag-correlation analysis will improve the ability to forecast demand. Sometimes several different times series must be united in a multiple correlation. In particular, when the inputs to the demand generator which we control can be related to the demands generated, both correlational techniques and functional analysis will provide a basis for forecasting. They also grant an element of control over the course of future events.

Some form of stability must exist for any of the methods above to work. The stability of sequence or time-ordering may exist and yet not meet the Shewhart criterion. Consequently, it is of the utmost importance to discern the nature of any stable pattern which the demand generator seems to follow. Based on what we learn, the correct inventory model can be chosen. Without such knowledge, the development of inventory decision rules becomes an exercise in decision-making under uncertainty. So again we see that only the systemic point of view makes sense when an inventory study is undertaken. Each of the parts are interlinked with each other and a decision in one area produces an effect on all the others.

Information theory provides us with one means for determining whether

* An intensive treatment of the implications of various methods is given in R. G. Brown, *Statistical Forecasting for Inventory Control* (New York: McGraw-Hill Book Co., Inc., 1959).

or not a regular pattern exists in the demand generated over a period of time. The formulas of information theory provide two different measures which are useful in this regard.

(1) H is a measure of the information content of a demand series. Stated in another way, it is the amount of variety which is an inherent characteristic of the demand generator. It also reflects the precision with which we observe and record demand. Everything else being equal, a system which lumps demand into the closest dozen has a lower H measure than one which treats each unit as an individual. As the variance of the demand distribution increases, the H value increases. No matter how large or small the demand range, the maximum entropy, $H(\text{maximum})$, occurs when all possible demands falling within that range are equally likely.

(2) R is a relative measure of the degree of ordering for any particular demand range. $R = 0$, when the system has no underlying pattern. $R = 1$, when the system is totally ordered and therefore totally predictable.

The formulas* are direct:

$$H(z) = - \sum_1^j p_j \log p_j$$

and

$$R(z) = 1 - \frac{H(\text{measured})}{H(\text{maximum})}$$

where j is the number of classes into which the demand range is divided; $H(z)$ is the average measure of the variety of demand, z ; $R(z)$ is the relative measure of complexity of the demand distribution; and p_j is the probability that the j th demand state will occur. If the R measure is large it indicates that only a few of the total possible number of demand states predominate. On the other hand, if R is low, then the implication exists that effective forecasting will be difficult.

We now observe digram or one-step transitional probabilities. Thus, we find the set of probabilities, p_{ij} which is the probability that the i th demand state is followed by the j th demand state. The steady state of the set of transition probabilities is obtained when we set $u'_j = u_j$, from

$$u'_j = \sum_i p_{ij} u_i \quad (\text{for all } j)$$

$$\sum_j p_{ij} = 1.00$$

and

$$\sum_j u_j = 1.00$$

It is then a simple matter to obtain the one-stage dependency measures of both H and R . The H measures are derived for each demand state j . Thus

$$H(z = j) = - \sum_i p_{ji} \log p_{ji}$$

Then

$$H(z)' = \sum_j u_j H(z = j)$$

* One of the major contributors to the theory of information is C. Shannon whose original work in this field was reported in Shannon and Weaver, *The Mathematical Theory of Communications* (Urbana, Ill.: The University of Illinois Press, 1949).

where $H(z)'$ is the measure of the average variety considering the one-step transition. The value of $H(\text{maximum})$ remains unchanged and we therefore obtain $R(z)'$ which is the redundancy based on the one-step dependency. If $R(z)' > R(z)$, then we know that an element of sequence stability is present. Other kinds of dependencies can be tested as well (two and more steps and various combinations). In this way, if stability of any kind exists it can be detected and forecasts can then be improved by incorporating whatever knowledge has been gained.

Another method for determining sequence patterns is by means of auto-correlation. The equation which is normally used is:

$$R(\tau) = \frac{1}{T} \int_0^T z(t)z(t - \tau) dt$$

where T equals the number of observations made, $z(t)$ is the difference between the expected demand and the actual demand at time t , $z(t - \tau)$ is the difference between the expected demand and the actual demand at time $(t - \tau)$. This auto-correlation function produces a series of numbers, one for each value of τ that is used. Thus, when $\tau = 0$, we are squaring $(z - \bar{z})$. Dividing by T produces the variance of the demand distribution. If a real basis exists for the historical forecast, the same result would be obtained when $\tau = 12$. That is, with a twelve month lag, the demand series is identical. Then we would have an auto-correlation coefficient of one. The auto-correlation coefficient is simply the auto-correlation function divided by the variance, or

$$\frac{R(\tau)}{R(0)}$$

Using this approach, stability in demand sequences can also be discovered and appropriate action can be taken to assure the proper functioning of the inventory system.

MULTIPLE DISTRIBUTIONS

When various distributions such as demand and lead time co-exist, it is desirable to ascertain the degree of dependency that exists between them. For this purpose information theory can be used again. Assume that $\{z\}$ is the demand distribution with i states and $\{y\}$ is the lead time distribution with j states. Then,

$$H(z, y) = - \sum_i \sum_j p(i, j) \log p(i, j)$$

where $p(i, j)$ is the joint probability of occurrence of the i th state of z and the j th state of y . $H(z, y)$ is the joint variety of these two different kinds of events. If dependencies exist, then:

$$H(y|z) = - \sum_i \sum_j p(i, j) \log p(j|i)$$

where $p(j|i)$ is the conditional probability that j will occur if an i has already occurred. This is the conditional variety of y given z .

It can be shown that:

$$H(z, y) = H(z) + H(y|z)$$

$H(z)$ is as previously defined. When z and y are independent, then:

$$H(z, y) = H(z) + H(y)$$

This case is equivalent to $H(\text{maximum})$ and so we can write for R :

$$R(z, y) = 1 - \frac{H(y|z)}{H(y)} \quad \text{where } H(y|z) \leq H(y)$$

This measure describes the degree of dependency of the elements in our joint distribution. The value it takes will influence the method we use to make predictions of demand in the lead time. Of course, it applies as well to joint distributions other than demand and lead time.

First of all, let us note that if dependencies exist, a joint distribution must be used for simulations. We cannot independently generate the separate distributions since they are in fact linked.

If independence does not exist then it would appear to be necessary to make forecasts conditional on the length of the lead time interval—and then to forecast the lead time interval. The procedure would be cumbersome and can be avoided by means of the forecast-error distribution. We obtain a measure† $F(z)(L.T./F.P.)$, where F.P. is the period for which the demand is forecast. F.P. can be quarterly, semiannually, etc., depending upon the practices of the company and the characteristics of the demand. Let us call the actual demand in the j th lead time period, z_j^* . Then we can compare for as many periods as we have data:

$$z_j^* - F(z) \left(\frac{L.T._j^*}{F.P.} \right) = \varepsilon$$

$L.T._j^*$ is the actual lead time experienced for the j th period. When a sufficient number of ε values have been obtained we can construct the distribution $\{\varepsilon\}$ with mean $\bar{\varepsilon}$ and variance σ_ε^2 . For the Q -model the reorder point will then be:

$$\frac{F(z)F(L.T.)}{F.P.} + \bar{\varepsilon}_{L.T.} + K\sigma_{\varepsilon_{L.T.}}$$

where $F(L.T.)$ is the forecasted lead time.

For the P -model the error must be estimated for the lead time plus the review period. The M -level would then be:

$$[F(L.T.) + R.P.] \left[\frac{F(z)}{F.P.} \right] + \bar{\varepsilon}_{R.P. + L.T.} + K\sigma_{\varepsilon_{R.P. + L.T.}}$$

† For a more detailed discussion of the implications of this general approach, see John Fox, *An Analysis of Inventory Management* (New York: Management Services, Remington Rand, UNIVAC, Division of Sperry Rand Corporation, 1957)

The effect of the dependencies will be to cause larger values of $\bar{\epsilon}$ and σ_{ϵ} than should otherwise occur. If the distributions are independent of each other then simulation can be easily achieved and in addition the variance of the error distribution can be modified in accordance with changes in the time period that is to be described; i.e.,

$$\left(\frac{a+b}{a}\right) \sigma_a^2 = \sigma_{a+b}^2$$

But above all, it must be remembered that if constant changes occur in the forecast, then dynamic models are indicated and the usual inventory models will not provide an optimal inventory system.

UPDATING

We have shown that delays can be very costly. Transmission and data-processing delays present a false picture of the stock on hand. The result can produce weird behavior and abnormally high costs. Similarly delays in detecting changes in the input forms and adapting to these can cripple the inventory system. Updating is the procedure for recognizing changes and incorporating these changes in the decision rules. Not only must demand and lead time be updated, but ordering costs, carrying costs, prices paid for stock, discounts available, forecasting errors, and service levels are also subject to change.

The optimal interval for updating is in itself a question of balancing marginal costs. The total cost of updating decreases as the interval for updating increases. What is the counter-balancing cost? It is the entire set of systemic costs associated with delays. We have tried to define a number of these as we proceeded. Of course, there are others, but one must compromise by studying only those which seem to be most relevant. In any case, it is likely that the delay costs go up exponentially. If an updating period has been chosen, it is possible to impute a cost for the delay. Thus,

$$T.C_T = \delta_T t^k + \frac{UT}{t}$$

where $T.C_T$ = total cost for the period T ,

δ_T = the cost per total time period for delay,

k = constant,

U = the cost per updating cycle,

t = the number of time periods in an updating cycle.

Taking the derivative and setting the resulting expression equal to zero.

$$\frac{\partial T.C_T}{\partial t} = \delta_T t^{k-1} - \frac{UT}{t^2} = 0$$

Whence,

$$t_o = \left(\frac{UT}{\delta_T}\right)^{1/(k+1)}$$

We see that if k is larger than one, the optimal updating cycle will be quite small. Further, if a computer is used, then U will be relatively small and

again a short cycle is indicated. If t is chosen by intuition, then the imputed cost of delay will be:

$$\delta_T = \frac{UT}{t^{k+1}}$$

At the least this information can be given to the inventory manager.

Even as simple a parameter as price can significantly disturb the performance of the system if it is inaccurately estimated. Carrying costs of both normal usage and reserve stocks will reflect the accounting system that is being used; e.g., FIFO, LIFO, etc. Although the ordering cost savings are relatively insensitive to poor estimates the reserve stock costs and the forecast error costs are not so unresponsive. The total sum of delay costs must therefore be balanced against the updating costs.

In order to derive the updating costs it is again necessary to embark upon the patient exploration of detail. We will list here the steps which are required to prepare the appropriate flow diagram for one possible updating routine where primes represent updated estimates:

- (1) Information which is in memory, or filed:
 - (a) Expected annual demand = $E[z]$.
 - (b) Reorder Point = P .
 - (c) Expected forecasting error = $E[\varepsilon]$.
 - (d) Variance of forecasting error = σ_ε^2 .
 - (e) Expected lead time = $E[L.T.]$.
 - (f) Variance of lead time = $\sigma_{L.T.}^2$.
 - (g) Average price paid per unit for all units, in stock = $\bar{c}$.
(Calculated on a first-in-first-out basis, FIFO.)
 - (h) Cost per order = C_r .
 - (i) Annual carrying cost rate = C_c .
 - (j) Specified service factor = α .
 - (k) All information supporting the above parameters.
- (2) Updating routine, x_0 :
 - (a) Executive fiat: Change C_r , C_c , if instructed (C'_r , C'_c).
 - (b) Using first-in-first-out (FIFO), subtract units issued at price paid. Add units received at price paid. Compute new $\bar{c}'$.
 - (c) Input = Quantity issued in the month = $Q.I.$
 - (d) Compute new $E[z']$ by whatever forecasting method is used.
 - (e) Compute new $x_0 = (2E[z']C'_r/\bar{c}'C'_c)^{1/2}$.
- (3) Updating routine; reorder point:
 - (a) Input = lead times experienced.
 - (b) Modify $E[L.T.]$ and $\sigma_{L.T.}$, ($E[L.T.]'$, $\sigma'_{L.T.}$).
 - (c) Compute new reasonable maximum lead time = $L.T'_{\max}$. That is, new $E[L.T.]' + K_1\sigma_{L.T.}$ where K_1 is the protection level specified for $L.T.$ by α . ($L.T_{\max}$ is expressed as a fraction of a year.)
 - (d) Obtain expected demand in the period $L.T_{\max}$, e.g., new $E[z]' \times L.T'_{\max}$.

- (e) Compare last period's forecast $E[z]$ against Q.I. and use error to obtain new $E[\varepsilon]$ and σ_ε , $(E[\varepsilon]', \sigma_\varepsilon')$
- (f) If new $E[\varepsilon]'$ is large, go back to step 2(c) and investigate forecasting method. If new $E[\varepsilon]'$ is reasonable, move to next step.
- (g) Obtain the variance of the error distribution in the period $L.T_{\max}$. This is,

$$\sigma_\varepsilon'[L.T_{\max}] = \sigma_\varepsilon' \sqrt{L.T_{\max}'}$$

- (h) Compute new reorder point, where K_2 is the protection level specified for reserve stock.

$$P' = E[z]' \cdot L.T_{\max}' + K_2' \sigma_\varepsilon' \sqrt{L.T_{\max}'}$$

- (i) Compare new reserve stock with old reserve stock. Thus,

$$(K_2' \sigma_\varepsilon' \sqrt{L.T_{\max}'} - K_2 \sigma_\varepsilon \sqrt{L.T_{\max}}) = q$$

If $q = 0$, go to next item to be updated. If $q > 0$, order an amount q with next order—then go to next item. If $q < 0$ decrease next order by that amount—then go to next item.

Figure 9-25 presents these steps in flow form. Still further detail is required before the actual routine can be used. For example, consider the further specification of $\bar{c}$ updated to $\bar{c}'$. This involves the following computation:

$$\sum_{j=1}^{j=n} \frac{c_j}{n} = \bar{c}$$

where n equals the total number of items in stock and c_j equals the price of the j th item. To accomplish this updating we have:

$$\bar{c}' = \frac{\sum_{j=1}^{j=n} c_j - \sum_{j=1}^{j=k} c_j + \sum_{j=n+1}^{j=m} c_j}{n - k + m - n} = \sum_{j=k+1}^{j=m} \frac{c_j}{m - k}$$

where k = the number of units issued, and $(m - n)$ = the number of units received. Set $k = 0$, which gives us the updated

$$\bar{c} = \sum_{j=1}^{j=m} \frac{c_j}{m}$$

81. CONCLUSION

We have tried to show that an inventory study is not merely a matter of finding a model which expresses a particular management policy. Certainly, the study is not completed until the operational characteristics of the system have been determined and taken into account. It has been our intention to emphasize repeatedly that the choice of decisions rules is dependent upon the multitude of systemic factors which operate, whether or not we acknowledge

PROBLEMS

1. The demand for paper towels in an hotel is approximately 2000 rolls per day. The towels are issued from the stockroom in lots of 10. The hotel is weighing the advantages of perpetual inventory control as compared to periodic reordering on a fixed cycle. It is determined that each issue transaction would cost \$.02 under the perpetual inventory system and \$.005 under the fixed cycle system. Further the variable order cost component for reordering the towels from the vendor is \$1.50 for the perpetual inventory system and \$1.75 for the fixed order cycle system. The cost per roll is \$.08. Management estimates C_c at .24 per year. Also, outages are not allowed to exceed one time per year. Delivery takes two days and the vendor has never deviated from this interval.

- (a) Assume a constant daily demand. Which system should be chosen?
- (b) If the demand is horizontally distributed with upper and lower bounds of 1500 and 2500, which system should be chosen?

2. The towels discussed in the example above constitute $\frac{1}{3}$ of the hotel's inventory data-processing load. One man whose salary is \$8000 per year handles this load. Is there a more efficient ordering policy that can be followed? (Give your answers for both the constant demand and the horizontal distribution.)

3. We will change the conditions stated in 2. Only $\frac{1}{8}$ of the employee's salary of \$8000 is allocated to the processing of towels inventories. What should be done for both types of distributions?

4. Derive the $k = 10$, Erlang service distribution and then convert the cumulative distribution to the appropriate density function. Comment on the relationship of the shape you have derived to the various data-processing steps which are likely to be encountered in practice.

5. Determine the response of the regulator P -model to a sinusoidal input by means of simulation.

6. Simulate the alternative flow pattern in Fig. 9-12 (i.e., $A \rightarrow B \rightarrow C \rightarrow D \rightarrow E \rightarrow CW \rightarrow A$). What is the effect of transmission through the additional linkages? What steps might be taken to improve regulation?

7. Using the delay formulas:

$$P(z) = D(z)z^{k+1}/(1 + z^{k+1} - z)$$

and

$$S(z) = D(z)/(1 + z^{k+1} - z)$$

determine the effect of a two-period delay; i.e., $k = 2$. Comment on your results.

8. Show that:

$$R = 1 - \frac{H(y|z)}{H(y)} = 1 - \frac{H(z|y)}{H(z)}$$

Discuss the use of this relationship.

9. We obtain the following two years of demand data:

n	1	2	3	4	5	6	7	8	9	10	11	12
z_n	1	3	5	6	5	7	5	3	2	4	3	3
z_{2n}	2	1	3	5	6	8	7	5	4	2	4	2

- (a) Derive the auto-correlation coefficient for $\tau = 1, 2$, and 3.
(b) Compare the result above when $\tau = 1$ with the use of information theory and digram transitions.

appendix

SECTION 25

The various inequalities we have derived take the form

$$P(|y - \bar{z}| \geq k\theta) \leq g(k, \theta)$$

where θ designates any of the multiples of k which we used and $g(k, \theta)$ is some function of k and θ . From this expression we can derive the two inequalities

$$F(\bar{z} - k\theta) \leq g(k, \theta)$$

and

$$F(\bar{z} + k\theta) \geq 1 - g(k, \theta)$$

We would like to use these inequalities together with the fact that

$$F(x) = \frac{C_u - c}{c}$$

is the minimizing expression for

$$E.F_x = cx + C_u \int_x^{\infty} (y - x)f(y) dy$$

as a way to achieve a distribution-free analysis of this kind of inventory problem.

Unfortunately, there are at least three difficulties which prevent this approach. First, if we proceed thus we have no expression for the expected cost itself. If we did succeed in finding the minimizing value for x we would

be unable to determine what the actual expected cost for the determined x would be. Second, on which side of the mean should we look for the minimum? The second derivative of $E.F_x$ is $cf(x)$ and this is always greater than or equal to zero for all x . This means that the minimum of $E.F_x$ can occur anywhere so we have no way of knowing on which side of the mean it occurs (since we have no expression for $E.F_x$). The third problem is more than a difficulty—it is fatal. This is that there is no logical basis for this procedure. The procedure we followed in Sections 23 and 24 was to write a maximum possible value for the expected cost. We then minimized this maximum value over all order quantities. This constitutes a perfectly reasonable utilization of the general minimax approach. But the approach we are now considering has no such logical justification. The distinction is important enough to justify an illustration of the difference. For simplicity let us suppose that we have to consider not all possible probability distributions but only a set of one-parameter distributions. In other words, for each value of the parameter, q , there is a specific distribution. This distribution is uniquely and completely characterized by the value of q . The Poisson distribution is an example of such a class of one-parameter distributions. We will assume that an inventory problem must be resolved and that we know that the demand distribution will be one of this class of distributions but we don't know which one. In other words, we are simplifying the general problem by limiting the possible distributions to one type. Under these circumstances our expected cost equation takes the form: $E.F_x = g(x, q)$. This equation simply states that the expected cost is a function of the order size, x , and the actual parameter of the demand distribution, q . Our previous procedure, then, was to reason as follows:

The maximum possible expected cost is determined by

$$\frac{\partial E.F_x}{\partial q} = 0$$

This gives a specific value of q , say q_0 . Then $g(x, q_0)$ is the maximum possible expected cost and this is minimized over x by

$$\frac{\partial g(x, q_0)}{\partial x} = 0$$

This gives a minimizing value of x , say x_0 . The minimax solution is, then, $E.F_x = g(x_0, q_0)$. Now, we might consider the reverse, or maximin, approach:

First, minimize $E.F_x$ with respect to x by

$$\frac{\partial E.F_x}{\partial x} = 0$$

and obtain the minimizing value of x , say x_1 . Then maximize $g(x_1, q)$ with regard to q by

$$\frac{\partial g(x_1, q)}{\partial q} = 0$$

and obtain the maximizing value of q , say q_1 . Then $E.F_x = g(x_1, q_1)$ is the maximin solution. We could also consider the conditions under which

$$g(x_1, q_1) = g(x_0, q_0)$$

However, neither of these is the procedure we are now considering. This would go as follows:

First, minimize $E.F_x$ with respect to x by

$$\frac{\partial E.F_x}{\partial x} = 0$$

and obtain the minimizing value of x , say x_1 . Let $g'(x, q)$ be the derivative of $g(x, q)$ with respect to x . Now maximize $g'(x_1, q)$ with respect to q by

$$\frac{\partial g'(x_1, q)}{\partial q} = 0$$

giving the maximizing value of q , say q_2 . This solution, whatever it might be called, would be $E.F_x = g(x_1, q_2)$. Clearly, this suggested procedure cannot be justified. What we actually get is the minimizing value of x and the value of q which maximizes the rate of change of $E.F_x$ for that value of x .

We must, therefore, use a different approach. Let us assume that the amount to be ordered, x , will be greater than the mean of the demand distribution. If we are going to use one of our inequalities we can write $x = \bar{z} + k\theta$, where k is to be determined. With this expression for x we can write the expected cost equation as

$$\begin{aligned} E.F_x &= c(\bar{z} + k\theta) + C_u \int_{\bar{z} + k\theta}^{\infty} (y - \bar{z} - k\theta)f(y) dy \\ &= c(\bar{z} + k\theta) + C_u \left[\int_{\bar{z} + k\theta}^{\infty} f(y) dy - (\bar{z} + k\theta) \int_{\bar{z} + k\theta}^{\infty} f(y) dy \right] \end{aligned}$$

Let us consider the first integral in this expression:

$$\begin{aligned} \int_{\bar{z} + k\theta}^{\infty} yf(y) dy &= \int_{\bar{z} + k\theta}^{\bar{z} + (k+1)\theta} yf(y) dy + \int_{\bar{z} + (k+1)\theta}^{\bar{z} + (k+2)\theta} yf(y) dy + \int_{\bar{z} + (k+2)\theta}^{\bar{z} + (k+3)\theta} yf(y) dy + \dots \\ &\leq [\bar{z} + (k+1)\theta] \int_{\bar{z} + k\theta}^{\bar{z} + (k+1)\theta} f(y) dy + [\bar{z} + (k+2)\theta] \int_{\bar{z} + (k+1)\theta}^{\bar{z} + (k+2)\theta} f(y) dy + \dots \\ &\leq [\bar{z} + (k+1)\theta] \left[\int_{\bar{z} + k\theta}^{\infty} f(y) dy - \int_{\bar{z} + (k+1)\theta}^{\infty} f(y) dy \right] \\ &\quad + [\bar{z} + (k+2)\theta] \left[\int_{\bar{z} + (k+1)\theta}^{\infty} f(y) dy - \int_{\bar{z} + (k+2)\theta}^{\infty} f(y) dy \right] + \dots \\ &\leq (\bar{z} + k\theta) \int_{\bar{z} + k\theta}^{\infty} f(y) dy + \dots \end{aligned}$$

$$\dots + \theta \left[\int_{\bar{z} + k\theta}^{\infty} f(y) dy + \int_{\bar{z} + (k+1)\theta}^{\infty} f(y) dy + \int_{\bar{z} + (k+2)\theta}^{\infty} f(y) dy \right]^* + \dots$$

But from our inequality we have that

$$P[(y - \bar{z}) \geq (k + i)\theta] \leq P[|y - \bar{z}| \geq (k + i)\theta] \leq g(k + i; \theta)$$

therefore,

$$\int_{\bar{z} + k\theta}^{\infty} yf(y) dy \leq (\bar{z} + k\theta) \int_{\bar{z} + k\theta}^{\infty} f(y) dy + \theta \sum_{i=0}^{\infty} g(k + i; \theta)$$

Returning to our expression for $E.F_x$, we can write

$$E.F_x \leq c(\bar{z} + k\theta) + C_u \theta \sum_{i=0}^{\infty} g(k + i; \theta)$$

The worst possible case would be when the equality holds.

We would now like to sum the infinite series on the right-hand side of this equation. Given a specific inequality we can use some such device as the Euler-Maclaurin formula for this purpose.[†] For example, if we use the Tchebycheff inequality we have:

$$\theta = s$$

$$g(k + i; \theta) = \frac{1}{(k + i)^2}$$

The Euler-Maclaurin formula gives the sum of the resulting series as

$$\sum_{i=0}^{\infty} \frac{1}{(k + i)^2} = \frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} - \frac{1}{30k^5} + \frac{1}{42k^7}$$

to a sufficient degree of accuracy for any ordinary purpose. Other inequalities can be similarly treated.

If we assume that the optimal order size is less than the average demand we can write $x = \bar{z} - k$ and proceed as before. We will consider here only the case where we are using the Tchebycheff inequality. For this case we find that

$$E.F_{\bar{z} - ks} = E.F_{(\bar{z} + ks)} + C_u \left[\int_{\bar{z} - ks}^{\bar{z} + ks} yf(y) dy - (\bar{z} - ks) \int_{\bar{z} - ks}^{\bar{z} + ks} f(y) dy \right]$$

It is this extra term which gives us trouble in establishing good approximations in this case. We can immediately derive:

$$C_u \left[\int_{\bar{z} - ks}^{\bar{z} + ks} yf(y) dy - \int_{\bar{z} - ks}^{\bar{z} + ks} f(y) dy \right] \leq 2ks C_u \int_{\bar{z} - ks}^{\bar{z} + ks} f(y) dy$$

Unfortunately, our inequality gives only a lower bound here:

$$\int_{\bar{z} - ks}^{\bar{z} + ks} f(y) dy \geq 1 - \frac{1}{k^2}$$

* We have not given any justification for these various steps. Since infinite series are involved we must observe suitable precautions. We will be satisfied here to note simply that the steps taken are all correct if the demand distribution has a finite upper limit, which, of course, is true of all practical demand distributions.

† See Edmund Whittaker and G. Robinson, *The Calculus of Observations* (London: Blackie and Son, Ltd., 1958), Chap. 7.

We need an upper bound. There is an obvious gross upper limit of one. Using this we can write

$$E.F_z \leq c(\bar{z} - ks) + 2ksCu + Cus \left(\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right)$$

If we assume a symmetrical unimodal distribution in this case we can develop the following inequality

$$E.F_z \leq c(\bar{z} - ks) + C_u ks + \frac{2}{9} C_u S \left[\frac{1}{k} + \frac{1}{2k^2} + \frac{1}{6k^3} \right]$$

SECTION 39

The expression for the convolutions of our demand distribution is

$$f\left(\frac{z}{s} + t'\right)(y) dy = \frac{\sqrt{z}}{\sqrt{x + t'z} s \sqrt{2\pi}} \exp\left(\frac{-(y - \bar{z})^2 \sqrt{z}}{2(x + t'z)s^2}\right) dy.$$

Using this expression in the total cost equation we can minimize with respect to r , assuming a fixed x . This gives

$$f\left(\frac{r\sqrt{z}}{s\sqrt{x + t'z}}\right) = \frac{cC_c(x + t'z)^{3/2}s}{52Kz^{3/2}}$$

Note that multiplication by s_i has already been taken into account so the right-hand side of this expression can be looked up directly as the ordinate of the normal distribution in standard form.

This is a particularly unpleasant equation to work with. Taking the derivative of the T.C. equation with respect to x gives

$$\frac{cC_c}{104\bar{z}} \left[1 + \frac{r}{x + t'z} \right] = \frac{C_r}{x^2} + \frac{K}{(x + \bar{z}t')^2} \left[1 - F\left(\frac{r\sqrt{z}}{s\sqrt{x + t'z}}\right) \right]$$

We can start with the value of x obtained from the optimal lot-size equation and solve for r by means of our first equation. Then we can use this value of r and find the value of x which satisfies our second equation. Continuing in this fashion we can converge quite rapidly on the two values which simultaneously fulfill both equations.

We start with $x = 294.4$. We calculate

$$f\left(\frac{r\sqrt{50}}{5\sqrt{444.4}}\right) = f(.0671r) = .009094$$

$$r = 41.0$$

For this value of r we find $x = 284$. For this value of x we now find $r = 40.7$. This is close enough to the original value to be considered the solution within practical limits of accuracy.

SECTION 44

γ_1 is defined as m_3/s^3 . We have that

$$\mu_{3,i} = im_3 \quad \text{and} \quad s_i = \sqrt{i}s$$

Therefore,

$$\gamma_{1,i} = \frac{\mu_{3,i}}{s_i^3} = \frac{im_3}{is^3\sqrt{i}} = \frac{1}{\sqrt{i}} \frac{m_3}{s^3}$$

γ_2 is defined as $m_4/s^4 - 3$. We have that

$$\mu_{4,i} = im_4 + 3i(i-1)s^4 \quad \text{and} \quad s_i = s\sqrt{i}$$

Therefore,

$$\begin{aligned} \gamma_{2,i} &= \frac{\mu_{4,i}}{s_i^4} - 3 = \frac{im_4 + 3i(i-1)s^4}{i^2s^4} - 3 \\ &= \frac{m_4 + 3(i-1)s^4}{is^4} - 3 \\ &= \frac{m_4 - 3s^4}{is^4} \end{aligned}$$

but $m_4 = (\gamma_2 + 3)s^4$ so

$$\gamma_{2,i} = \frac{\gamma_2}{is^4}$$

In particular, when $\gamma_2 = 0$ then $\gamma_{2,i} = 0$ for all i .

SECTION 46

Since $s_i = s\sqrt{t+t'}$, we have the equation:

$$T.C. = \frac{52C_r}{t} + \frac{t\bar{z}cC_c}{2} + k\sqrt{t+t'}scC_c + \frac{52K}{(t+t')k^2}.$$

The necessary derivatives are

$$\begin{aligned} \frac{dT.C.}{dk} &= \sqrt{t+t'}scC_c - \frac{104K}{(t+t')k^3} = 0 \\ \frac{dT.C.}{dt} &= \frac{-52C_r}{t^2} + \frac{zcC_c}{2} + \frac{kscC_c}{2\sqrt{t+t'}} - \frac{52K}{(t+t')^2k^2} = 0. \end{aligned}$$

Solving the first of these for t gives $t = \frac{1}{k^2} \left(\frac{104K}{scC_c} \right)^{2/3} - t'$.

Substituting this expression for t in the second equation and solving for k gives

$$k = \left(\frac{104K}{scC_c} \right)^{1/3} / \sqrt{\sqrt{\frac{104C_r}{\bar{z}cC_c}} + t'}$$

Putting this in our expression for t gives

$$t = \sqrt{\frac{104C_r}{\bar{z}cC_c}}$$

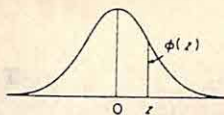
and, finally,

$$k = \sqrt[3]{\frac{104K}{(t+t')s_i c C_c}}.$$

TABLE OF RANDOM NUMBERS

96268	11860	83699	38631	90045	69696	48572	05917	51905	10052
03550	59144	59468	37984	77892	89766	86489	46619	50263	91136
22188	81205	99699	84260	19693	36701	43233	62719	53117	71153
63759	61429	14043	49095	84746	22018	19014	76781	61086	90216
55006	17765	15013	77707	54317	48862	53823	52905	70754	68212
81972	45644	12600	01951	72166	52682	97598	11955	73018	23528
06344	50136	33122	31794	86423	58037	36065	32190	31367	96007
92363	99784	94169	03652	80824	33407	40837	97749	18364	72666
96083	16943	89916	55159	62184	86208	09764	20244	88388	98675
92993	10747	08985	44999	36785	65035	65933	77378	92339	96454
95083	70292	50394	61044	65591	09774	16216	63561	59751	78771
77308	60721	96057	86031	83148	34970	30892	53489	41999	18021
11913	49624	28510	27311	61586	28576	43092	69971	44220	80410
70648	47484	05095	92335	55299	27161	64486	71307	85883	69610
92771	99203	37786	81142	44271	36433	31726	74879	89348	76886
78816	20975	13043	55921	82774	62745	48338	88348	61211	88074
79934	35392	56097	87613	94627	63622	08110	16611	88599	02890
64698	83376	87524	36897	17215	74339	69856	43622	22567	11518
44212	12995	03581	37618	94851	63020	65348	55857	91742	79508
82292	00204	00579	70630	37136	50922	83387	15014	51838	81760
08692	87237	87879	01629	72184	33853	95144	67943	19345	03469
67927	76855	50702	78555	97442	78809	40575	79714	06201	34576
62167	94213	52971	85974	68067	78814	40103	70759	92129	46716
45828	45441	74220	84157	23241	49332	23646	09390	13032	51569
01164	35307	26526	80335	58090	85871	07205	31749	40571	51755
29283	31581	04359	45538	41435	61103	32428	94042	39971	63678
19868	49978	81699	84904	50163	22625	07845	71308	00859	87984
14294	93587	55960	23149	07370	65065	06580	46285	07884	83928
77410	52195	29459	23032	83242	89938	40510	27252	55565	64714
36580	06921	35675	81645	60479	71035	99380	59759	42161	93440
07780	18093	31258	78156	07871	20369	53947	08534	39433	57216
07548	08454	36674	46255	80541	42903	37366	21164	97516	66181
22023	60448	69344	44260	90570	01632	21002	24413	04671	05665
20827	37210	57797	34660	32510	71558	78228	42304	77197	79168
47802	79270	48805	59480	88092	11441	96016	76091	51823	94442
76730	86591	18978	25479	77684	88439	35112	26052	57112	91653
26439	02903	20935	76297	15290	84688	74002	09467	41111	19194
32927	83426	07848	59327	44422	53372	27823	25417	27150	21750
51484	05286	77103	47284	05578	88774	15293	50740	07932	87633
45142	96804	92834	26886	70002	96643	36008	02239	93563	66429
12760	96106	89348	76127	17058	37181	74001	43869	28377	80923
15564	38648	02147	03894	97787	35234	44302	41672	12408	90168
71051	34941	55384	70709	11646	30269	60154	28276	48153	23122
42742	08817	82579	19505	26344	94116	86230	49139	32644	36545
59474	97752	77124	79579	65448	87700	54002	81411	57988	57437
12581	18211	61713	73962	87212	55624	85675	33961	63272	17587
00278	75089	20673	37438	92361	47941	62056	94104	45502	79159
59317	31861	62559	30925	23055	70922	47195	29827	68065	95409
59220	42448	70881	33687	53575	54599	69525	76424	98778	10459
00670	32157	15877	87120	13857	23979	38922	62421	03043	19602

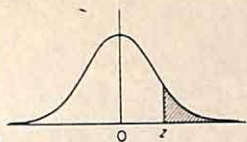
$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$



TO FOUR DECIMAL PLACES

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.3989	.3989	.3989	.3988	.3986	.3984	.3982	.3980	.3977	.3973
.1	.3970	.3965	.3961	.3956	.3951	.3945	.3939	.3932	.3925	.3918
.2	.3910	.3902	.3894	.3885	.3876	.3867	.3857	.3847	.3836	.3825
.3	.3814	.3802	.3790	.3778	.3765	.3752	.3739	.3725	.3712	.3697
.4	.3683	.3668	.3653	.3637	.3621	.3605	.3589	.3572	.3555	.3538
.5	.3521	.3503	.3485	.3467	.3448	.3429	.3410	.3391	.3372	.3352
.6	.3332	.3312	.3292	.3271	.3251	.3230	.3209	.3187	.3166	.3144
.7	.3123	.3101	.3079	.3056	.3034	.3011	.2989	.2966	.2943	.2920
.8	.2897	.2874	.2850	.2827	.2803	.2780	.2756	.2732	.2709	.2685
.9	.2661	.2637	.2613	.2589	.2565	.2541	.2516	.2492	.2468	.2444
1.0	.2420	.2396	.2371	.2347	.2323	.2299	.2275	.2251	.2227	.2203
1.1	.2179	.2155	.2131	.2107	.2083	.2059	.2036	.2012	.1989	.1965
1.2	.1942	.1919	.1895	.1872	.1849	.1826	.1804	.1781	.1758	.1736
1.3	.1714	.1691	.1669	.1647	.1626	.1604	.1582	.1561	.1539	.1518
1.4	.1497	.1476	.1456	.1435	.1415	.1394	.1374	.1354	.1334	.1315
1.5	.1295	.1276	.1257	.1238	.1219	.1200	.1182	.1163	.1145	.1127
1.6	.1109	.1092	.1074	.1057	.1040	.1023	.1006	.0989	.0973	.0957
1.7	.0940	.0925	.0909	.0893	.0878	.0863	.0848	.0833	.0818	.0804
1.8	.0790	.0775	.0761	.0748	.0734	.0721	.0707	.0694	.0681	.0669
1.9	.0656	.0644	.0632	.0620	.0608	.0596	.0584	.0573	.0562	.0551
2.0	.0540	.0529	.0519	.0508	.0498	.0488	.0478	.0468	.0459	.0449
2.1	.0440	.0431	.0422	.0413	.0404	.0396	.0387	.0379	.0371	.0363
2.2	.0355	.0347	.0339	.0332	.0325	.0317	.0310	.0303	.0297	.0290
2.3	.0283	.0277	.0270	.0264	.0258	.0252	.0246	.0241	.0235	.0229
2.4	.0224	.0219	.0213	.0208	.0203	.0198	.0194	.0189	.0184	.0180
2.5	.0175	.0171	.0167	.0163	.0158	.0154	.0151	.0147	.0143	.0139
2.6	.0136	.0132	.0129	.0126	.0122	.0119	.0116	.0113	.0110	.0107
2.7	.0104	.0101	.0099	.0096	.0093	.0091	.0088	.0086	.0084	.0081
2.8	.0079	.0077	.0075	.0073	.0071	.0069	.0067	.0065	.0063	.0061
2.9	.0060	.0058	.0056	.0055	.0053	.0051	.0050	.0048	.0047	.0046
3.0	.0044	.0043	.0042	.0040	.0039	.0038	.0037	.0036	.0035	.0034
3.1	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0025
3.2	.0024	.0023	.0022	.0022	.0021	.0020	.0020	.0019	.0018	.0018
3.3	.0017	.0017	.0016	.0016	.0015	.0015	.0014	.0014	.0013	.0013
3.4	.0012	.0012	.0012	.0011	.0011	.0010	.0010	.0010	.0009	.0009
3.5	.0009	.0008	.0008	.0008	.0008	.0007	.0007	.0007	.0007	.0006
3.6	.0006	.0006	.0006	.0005	.0005	.0005	.0005	.0005	.0005	.0004
3.7	.0004	.0004	.0004	.0004	.0004	.0004	.0003	.0003	.0003	.0003
3.8	.0003	.0003	.0003	.0003	.0003	.0002	.0002	.0002	.0002	.0002
3.9	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0001	.0001
4.0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09

$$A = \int_z^{\infty} \phi(z) dz$$



TO FOUR DECIMAL PLACES

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.8	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
3.9	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09

glossary

$A_0(t)$	= probability that an interval between arrivals is greater than t .
C_c	= carrying cost per unit of time.
C_o	= overstock cost per unit.
$C_r = F_r$	= cost per order.
$C_{rF} = G_1$	= fixed cost of an order.
C_s	= salvage value per unit.
C_u	= total cost of an understock of one unit.
c	= per unit cost of the item.
d	= discount.
$D(w_j)$	= actual demand in the j th interval of length w .
D_t	= input signal at time t .
E.O.C.	= expected opportunity cost.
E.F.	= expected payoff.
E	= fixed cost of the study.
F_{ij}	= payoff resulting from the i th strategy and j units of demand, e.g., F_{32} is the payoff resulting from the third strategy and a demand for 2 units.
$F(x)$	= cumulative probability distribution of demand.
$F(w_j)$	= forecasted demand in the j th interval of length w .
F.P.	= forecast period.
g_j	= the number of subgroups in class j .
H	= measure of information content (entropy).
$\bar{I}$	= total average dollar inventory.
ID	= the cost of idle facilities.

- K_n = opportunity cost of an understock of one unit.
 K_o = opportunity cost of an overstock of one unit.
 K = the number of standard deviations which are needed to produce a given confidence level.
 L.T. = lead time.
 L_q = the expected number of transactions which are waiting to be processed.
 L = the expected number of transactions which are waiting and being processed.
 M_0 = mode of the distribution.
 m'_r = moment of order r around x .
 m_r = moment of order r around the mean.
 M.C.N. = monte carlo numbers.
 P -model = fixed order period inventory model.
 P_t = output signal at time t .
 P_i = probability of being in state i .
 P = reorder point.
 $P(\quad)$ = probability of whatever is written in the parentheses.
 Q -model = perpetual inventory system.
 R = relative measure of degree of ordering (redundancy).
 $R(\tau)$ = autocorrelation function.
 $R(\tau)/R(0)$ = autocorrelation coefficient.
 r_n = per cent of total dollar demand of the n th item.
 R.P. = review period.
 r = reserve stock for the order period.
 s = standard deviation of a sample.
 λ = roots of the ruin equation.
 S.O.H. = stock on hand.
 S.O.O. = stock on order.
 $S_0(t)$ = probability that service time is greater than t .

T_A	= expected interval between arrivals.
T_S	= expected servicing time.
T_{zc}	= total dollar demand.
t	= time between orders.
T.C.	= total cost.
$T.C_{opt}$	= optimal total cost.
T.O.	= total orders.
T.I.	= total average inventory.
$T.C_i$	= total cost of the i th item.
U	= the cost per updating cycle.
u_f	= the probability of ruin.
V_r	= the variable cost per order.
v_D	= cost of estimating demand characteristics per estimate.
v_c	= cost of estimating cost factors per estimate.
v	= the average number of units per issue transaction.
W_q	= the expected waiting time of a transaction before it begins to be processed.
W	= the expected total time that a transaction spends in the system, waiting and being processed.
x	= amount to order.
x_o	= optimal amount to order.
z	= demand per unit of time.
z	= expected demand per unit of time.
$zc(n)$	= dollar demand for the n th item.
α	= service level = $1 - F(x)$.
β	= variance of demand usage distribution.
γ_1	= measure of skewness.
γ_2	= measure of kurtosis.
δ	= delay.
δ_T	= the cost per total time period for delay.

- $\varepsilon(w_j)$ = error term which corresponds to the j th interval of length w .
 ε_t = error measured at time t .
 ζ = the company's available capital as a multiple of study cost.
 κ = available capital.
 λ = Lagrangian multiplier.
 λ'_r = absolute moment of order r around x .
 λ_r = absolute moment of order r around the mean.
 μ_r = moment of order r .
 μ_1 = mean of the distribution.
 ρ_n = present worth of \$1.00 obtained at the end of each year for n years.
 $\sigma^2 = m_2$ = variance of the distribution.
 ϕ = expected percent savings.
 χ = study costs.
 ψ = savings per year.
 $\mathcal{Z}$ = transform variable for the generating function.
 $|y - x|$ = absolute value of $(y - x)$
 $>$ = is greater than
 $<$ = is less than
 $\geq$ = equal to or greater than
 $\leq$ = equal to or less than.
 $f(y)$ = function of demand.
 $f^i(y) dy$ = probability distribution of demand convoluted i times.
 $\sum_i$ = operator of addition, where i is the dummy of summation.
 $\int_a^b$ = operator of integration, where lower limit is a , and upper limit is b .
 $\frac{dy}{dx}$ = derivative of y with respect to x .
 $\frac{\partial y}{\partial x}$ = partial derivative of y with respect to x .

bibliography

CYBERNETICS

Ashby, W. R.: *An Introduction to Cybernetics*; John Wiley & Sons, Inc., New York, 1956.

de Latil, P.: *Thinking by Machine, A Study of Cybernetics*; (Trans. Y. M. Golla); Houghton Mifflin Co.; Boston, 1957.

DATA PROCESSING

Fox, J.: *An Analysis of Inventory Management*; Management Services, Remington Rand, UNIVAC, Division of Sperry Rand Corporation; New York, 1957.

Gregory, R. H., and R. L. Van Horn: *Automatic Data-Processing Systems*; Wadsworth Publishing Co., Inc.; San Francisco, 1960.

DYNAMIC PROGRAMMING

Bellman, R.: *Dynamic Programming*; Princeton University Press; Princeton, 1958.

Howard, R.: *Dynamic Programming and Markov Processes*; John Wiley & Sons, Inc.; New York, 1960.

Vazsonyi, A.: *Scientific Programming in Business and Industry*; John Wiley & Sons, Inc.; New York, 1958.

FORECASTING

Brown, R. G.: *Statistical Forecasting for Inventory Control*; McGraw-Hill Book Co., Inc.; New York, 1959.

INFORMATION THEORY

Pierce, J. R.: *Symbols, Signals and Noise*; Harper & Brothers; New York, 1961.

Shannon, C., and W. Weaver: *The Mathematical Theory of Communication*; The University of Illinois Press; Urban, Illinois, 1949.

MATHEMATICAL MODELS

- Arrow, K. T., S. Karlin, and H. Scarf: *Studies in the Mathematical Theory of Inventory and Production*; Stanford University Press; Stanford, California, 1958.
- Fetter, R. B., and W. C. Dalleck: *Decision Models for Inventory Management*; Richard D. Irwin, Inc.; Homewood, Illinois, 1961.
- Hanssmann, Fred: *Operations Research in Production and Inventory Control*; John Wiley & Sons, Inc., 1961.
- Holt, Modigliani, Muth and Simon: *Planning Production, Inventories and Work Force*; Prentice-Hall, Inc.; Englewood Cliffs, N. J., 1960.
- Massé, P.: *Les Reserves a la Regulation de l'Avenir dans la Vie Economique*, 2 vols. Hermann; Paris, 1946.
- Miller, D. W.: *The Utilization of Partial Information in the Quantitative Analysis of Decision Problems*; Unpublished Ph.D. Dissertation, Graduate School of Business; Columbia University; New York, 1962.
- Moran, P. A. P.: *The Theory of Storage*; Methuen & Co., Ltd.; London, and John Wiley & Sons, Inc.; New York, 1959.
- Whitin, T. M.: *The Theory of Inventory Management*; Princeton University Press; Princeton, N. J., 1959.

MOSTLY QUALITATIVE

- American Management Association: *Managing the Materials Function*; Report No. 35, Mfg. Division; AMA New York, 1959.
- Ammer, D. S.: *Materials Management*; Richard D. Irwin, Inc.; Homewood, Illinois, 1962.
- Magee, J. F.: *Production Planning and Inventory Control*; McGraw-Hill Book Co., Inc.; New York, 1959.

OPERATIONS RESEARCH

- Ackoff, R. L. (Ed.): *Progress in Operations Research*; John Wiley & Sons, Inc.; New York, 1961.
- Churchman, C. W., R. L. Ackoff, and E. L. Arnoff: *Introduction to Operations Research*; John Wiley & Sons, Inc.; New York, 1957.
- M. I. T.: *Notes on Operations Research, 1959*; The Technology Press; Cambridge, Massachusetts, 1959.

Saaty, T. L.: *Mathematical Methods of Operations Research*; McGraw-Hill Book Co., Inc.; New York, 1959.

PROBABILITY

Feller, William: *Probability Theory and Its Applications*; Vol. I, 2nd Edition; John Wiley & Sons, Inc.; New York, 1957.

Uspensky, J. V.: *Introduction to Mathematical Probability*; McGraw-Hill Book Co., Inc.; New York, 1937.

QUEUING THEORY

Cox, D. R., and W. L. Smith: *Queues*; Methuen & Co., Ltd.; London, and John Wiley & Sons, Inc.; New York, 1961.

Morse, P. M.: *Queues, Inventories and Maintenance*; John Wiley & Sons, Inc.; New York, 1958.

Operations Standards Division, Port of New York Authority: *A Collection of Delay Probability Formulas and Curves*; Port of New York Authority; New York, 1956.

Saaty, T. L.: *Elements of Queuing Theory*; McGraw-Hill Book Co., Inc.; New York, 1961.

SIMULATION

Alfandry-Alexander, Mark: *An Inquiry Into Some Models of Inventory Systems*; University of Pittsburgh Press; Pittsburgh, 1962.

Forrester, J. W.: *Industrial Dynamics*; John Wiley & Sons, Inc.; New York, 1961.

STATISTICS

Cramér, Harald: *Mathematical Methods of Statistics*; Princeton University Press; Princeton, 1946.

Schlaifer, R.: *Probability and Statistics for Business Decisions*; McGraw-Hill Book Co., Inc.; New York, 1959.

index

A

Absolute moments, 59
Accounting, 9 ff, 13, 14
cost, 94
Accuracy, level of, 197
Adaptive behavior, 297
Ad hoc methods, 13
Advertising, 316
Agricultural commodities, 17
Airline repair center, 242
Algorithm, 87, 136
Algorithm, square-root, 91
Amalgamation of orders, 13
Arithmetic mean, 24
Arrow, Kenneth, 17, 115, 116
Ashby, W. R., 289
Assembly diagrams, 216
Auditing, perpetual, 114
Autocorrelation, 322
Automata, 297
Automatic control, 220

B

Back-order, 12, 14
Bankruptcy, 16
probability of, 25
Banks, 14
Bayes, Thomas, 52
Behavior, adaptive, 297
Bellman, Richard E., 211
Binomial expansion, 118
Black box, 221
Boiler installation problem, 47
Borel, Emile, 17, 68
Bottlenecks, 283
Bounded rationality, 210, 257
Brand name, 316
Breadth of study, 175
Breakeven period, 173
Brown, R. G., 183, 320

Burden (*see* Overhead costs)
Buy or make (*see* Make or buy)

C

Cantelli, 68
Capacity and load, 266 ff
Carrying cost, 10, 35, 44
Cascade of decisions, 286, 288
Centralized ordering, 110
Central warehouses, 162 ff, 213
Certainty, 5, 6, 8, 17, 18
Characteristic functions, 119
Checks, clearance, 234
Christmas tree problem, 11, 19, 21 ff,
33, 39, 50, 318
Clark, John Maurice, 9
Clearance of checks, 234
Clerical processing, 9, 10
Coin-flipping, 55
Color-blindness, 57
Commodities, 143
Commodities, agricultural, 17
Complexity, increase, 176
Computer routine, 268
Computers, 18, 142, 158, 209
Concentration, 185
Conservatism, 51
Construction, 5, 20
materials, 78
Continuous problems, 40 ff
Continuous production, 10
Control, automatic, 220
Control monitor, 319
Control, reserve stock costs, 190 ff
Control systems, 13
Control systems, inventory, 113
Controls, updating, 316 ff
Control theory, 297
Convolutions, 126, 130, 153 ff, 160, 163,
188 f
Convolutions, moments, 135 ff

Corsage, moments, 153 ff
 Corsage problem, 47
 Cost accounting, 94
 Cost, carrying, 10, 35, 44
 Cost, delivery, 21
 Cost, deterioration, 11
 Cost, expediting, 12
 Cost, historical records, 14
 Cost, imputation, 15, 44 ff, 101, 156 f, 191
 Cost, installation, 13
 Cost, insurance, 11
 Cost, issue transaction, 261
 Cost, loss of goodwill, 12, 14, 45
 Cost, measurement, 13 f, 94 errors in, 14
 Cost, obsolescence, 11
 Cost of money, 10
 Cost, opportunity, 10, 26, 34, 42, 94
 Cost, opportunity, matrix, 26, 51
 Cost, opposing, 7, 8, 25, 40
 Cost, order, 260 ff
 Cost, out-of-stock, 12
 Cost, overhead, 9, 10
 Cost, overstock, 11, 20, 113
 Cost, pilferage, 11
 Cost, procurement, 9, 10
 Cost, reserve stock, control, 190 ff
 Cost, review, 106
 Cost, risk, 31 f
 Cost, set-up, 9, 10, 19
 Stock, stockage, 10 storage, 11
 Cost, sunk, 4, 9, 10
 Cost, systemic, 12 f, 18, 109, 110, 116, 124, 130, 171, 205, 295, 324
 Cost, understock 113 —
 Costs, 9 ff
 Costs, fixed and variable, 9, 14
 Costs, implementation, 13
 Costs, non-specific, 171
 Cost structures, other, 33 ff
 Cramer, Harold, 65, 70
 Critical damping, 312, 315
 Currency, inventory problem of, 3, 4
 Cybernetics, 288

D

Damping, critical, 312, 315
 Damping, over, 314 f

Data flow regulation, 283 ff
 Data-processing, 13, 18
 Data-processing, overloaded system, 272 ff
 Data-processing, periodic, 262 ff
 Data-processing, perpetual, 262 ff
 Data-processing routine, 263
 Data-processing study, 171
 Data-processing, underloaded system, 269 ff
 Decision criteria under uncertainty, 50 ff, 53 ff
 Decision criterion, 50
 Decision theory, 5, 49
 Defective items, 19
 Definitions, 3 f
 Degree of ordering, 321
 Delay, 308 ff
 Delivery costs, 21
 Demand, 5
 Demand distributions, moments, 59
 Demand generator, 218, 220, 221, 227, 243
 Demand probability distributions, 116 ff
 Demand, variance, 192
 Department stores, 4, 6, 11, 112
 Depth of study, 175
 Deterioration costs, 11
 Diminishing returns, in information, 60, 75
 Discounts, quantity, 84 ff
 Discrete problems, 36 ff
 Distribution-free analysis, 61, 62 ff, 66 ff, 73 ff
 Distributions:
 fixed demand, 6
 variable demand, 6
 Divergent oscillation, 312 ff
 Dollars, payoff measure, 16
 Dollar usage, variance of, 186
 Dorfman, Robert, 136
 Dvoretzky, Aryeh, 115
 Dynamic inventory problems, 6, 10, 11, 13, 18
 Dynamic programming, 87 ff, 136, 142 ff, 211, 227

E

Effectiveness, measures, 214 f
 Entropy, 321

index

Equilibrium probabilities, 133 ff
 Erland family, 275 ff
 Error rate vs. savings, 177
 Error regulation, 297
 Error sensor, 219
 Errors, cost measurement, 83 ff
 estimation, 175
 judgement, 176
 Errors, forecasting, 190 ff
 variance, 194 ff
 cost, 193 ff
 Euler-Maclaurin formula, 334
 Evaluation study, 171
 Evaluation study, cost of, 172
 Evaluation study, sample, 196 ff
 Exception, management by, 219
 Exhaustion, of study funds, 201 ff
 Expansion, binomial, 118
 Expected savings, 184, 185
 Expected value, 24
 Expediting costs, 12
 Exponential distribution, 149
 Exponential smoothing, 320

F

Fashion goods, 6, 11, 19
 Feedback, 298
 Feedback, inventory, 218
 Feedback loop, 219, 220, 234
 Feller, William, 62, 201, 202, 318
 Flat response systems, 316
 Flow charts, process, 216
 Flow diagram, master block, 215, 217
 Flow diagram of *Q*-Model simulation, 234 f
 Flow diagrams, 216, 234
 detailed, 233 ff
 receipt transaction, 286
 Flow, sequential, 216
 Fluctuating prices, 143
 Fluctuations, load, 274 ff
 Forecast, historical, 221, 230
 Forecasting, 20, 58, 316 ff
 errors, 190 ff
 cost, 193 ff
 variance, 194 ff
 Fox, John, 323
 Franklin, Philip, 42
 Frechet, Maurice, 68, 71
 Fromke, Robert, 298

Full and partial information:
 comparisons of, 74 ff
 Functional equation, 115
 Functions, characteristic, 119

G

Gambler's ruin, 201 ff
 Gantt charts, 216
 Gauss, Carl F., 70
 Generating functions, 118 f
 Generator, demand, 218, 220, 221, 227, 243
 Goal-oriented, 296
 Goodwill, 15
 cost, 14
 loss, 12, 45
 Gregory, 234
 Grouping, item, 181 ff, 189 f
 rules, 174
 Guidance system, 220
 Guided missiles, 210

H

Hanssman, Fred, 4
 Harris, Theodore E., 115
 Hedging, 18, 145 f
 Historical forecast, 221, 230
 Historical records of cost, 14
 Hogben, Lancelot, 52
 Holt, Charles C., 14
 Holzbock, W. G., 299
 Horizon, planning, 20, 88, 136 f
 Howard, Ronald, 136, 298
 Hurwicz Criterion, 54

I

Idle resources, inventory as, 4
 Implementation, 172
 costs, 13
 Imputation of costs, 15, 44 ff, 101, 156 f, 191
 Imputed value, 122
 Increase in complexity with size, 176
 Incremental analysis, 29 ff, 36, 39
 Industry, type of, 196 ff
 Inequality, Tchebycheff's, 62 ff
 Inertial systems, 319
 Information, age of, 289

Information, diminishing returns, 60, 75
 Information, full and partial, 74 ff
 Information, partial, 5, 8, 49 f, 152 f
 Information, perfect, 32
 Information theory, 320 ff
 Input-output models, 218 ff
 Installation costs, 13
 Insurance costs, 11
 Intangibles (*see* Tangibility of stock)
 Intangibles, storage, 14
 Intuitive and judgemental approach, 175 ff
 Inventory behavior under certainty, 80
 Inventory control systems, 113
 Inventory feedback system, 218
 Inventory models, 171
 Inventory study, 171
 Inventory study, elements of, 172 ff
 Issue transaction cost, 261
 Item grouping, 181 ff
 Item substitution, 12
 Items, grouping, 189 f
 Items, many, 93 ff
 effect, 211 ff
 simulation, 15, 16, 248 ff
 Items, multiple, 104 ff
 Items, number, 176 ff
 Items, sampling, 197 f

J

Jet engine replacements, 242
 Job-lot production, 10
 Joint distributions, 323
 Judgemental approach, intuitive and, 175 ff

K

Karlin, S., 17
 Kendall, Maurice G., 60
 Keynes, J. M., 17
 Keifer, J., 115
 Kurtosis, 60, 68, 154, 156

L

Labor load, 267
 Lagrangian multipliers, 95 ff

Lag-correlation analysis, 320
 Lags, variable, 146 ff
 Lag time (*see* Lead time)
 Laplace, P. S., 52
 Lead time, 5, 7, 20, 120, 122
 Lead-time generator, 236
 Leasing, 4
 Leibnitz's rule, 41 f
 Level of accuracy, 197
 Level of protection, 164
 Load and capacity, 266 ff
 Load fluctuations, 274 ff
 Load, labor, 267
 Loans, 11
 Logarithm, 16
 Lognormal distribution, 183
 Loop, feedback, 219, 220, 234
 Loss leader, 213
 Lot-size model, optimal, 79 ff
 Luce, R. Duncan, 51

M

Made-to-order items, 146
 Mail-order, 12
 Make or buy, 5, 6
 Management by exception, 219
 Man-machine charts, 216
 Many department effect, 214 ff
 Many item effect, 211 ff
 Many items, 93 ff
 Many item simulation, 15, 16, 248 ff
 Many location effect, 213 f
 Marginal principle, 75
 Market surveys, 32
 Markovian decision-making, 136 ff
 Markovian process, 131
 Marschak, Jacob, 115
 Masse, Pierre, 16
 Master block flow diagram, 215, 217
 Materials, construction, 78
 Materials management, block flow diagram, 284
 Matrices, switching, 130 ff
 Maximin criterion, 51
 Mean, arithmetic, 24
 Measuring costs, 13 ff, 94
 errors in, 14
 Measures of effectiveness (moe), 214 ff
 Mechanized equipment, 297
 Miller, David W., 16, 72, 201

Minima, 41
 Minimax criterion, 51
 Missiles, guided, 210
 Modigliani, Franco, 14
 Moe, 217, 219, 222
 Moments, absolute, 59
 Moments around the mean, 59
 Moments of convolutions, 153 ff
 Moments of demand distribution, 59
 Monitor, 219
 Monitor, control, 319
 Money, cost of, 10
 Monte Carlo, 198
 Monte Carlo simulation, *Q*-model,
 227 ff
 Monthly ordering policy, 94
 Morse, Philip M., 279
 Multiple items from one supplier,
 104 ff, 124
 Multiple orders, 78
 Multiple warehouses, 7, 110 ff
 Muth, John F., 14

N

New products, 5, 14
 Newsstand dealer, 19
 Nomograms, 160, 162
 Non-parametric methods, 61
 Non-specific costs, 171
 Nonstable distributions, 319 ff
 Nonstationary distributions, 319
 Number of items carried, 176 ff

O

Objectives of carrying inventory, 8,
 15 ff
 Objectivist school, 55
 Obsolescence costs, 11
 One-shot decisions, 6, 7
 On-line, 234
 Operating system, satisfactory, 260
 Operator charts, 216
 Operators, 37
 Opportunity cost matrix, 26, 51
 Opportunity costs, 10, 26, 34, 42, 94
 Opposing costs, 7, 8, 25, 40
 Optimality, 9, 14, 15
 principle of, 214

Optimal lot-size formula, 126
 Optimal lot-size model, 79 ff, 93
 varying demand, 86 ff
 Optimal policy curve, 93 ff, 103
 Optimal selling policy, 142 ff
 Orchid problem, 47
 Order cost, 9, 260 ff
 Ordering, centralized, 110
 Ordering, degree, 321
 Orders, amalgamation, 13
 Orders, multiple, 78
 Ordering, monthly, 94
 "Orphan" machine tools, 33
 Oscillation, divergent, 212 f
 Oscillation, undamped, 314
 Out-of-stock cost, 12
 Overdamping, 314 f
 Overhead costs, 9, 10
 Overloaded data-processing systems,
 272 ff
 Overstock cost, 11, 20, 113
 Overtime, 46

P

Partial information, 5, 8, 49 f, 58 ff,
 152 f
 Partial means, 72
 Payoff matrix, 21 ff
 Payoff measure, dollars, 16
 Peakedness, 60
 Pearson, Karl, 60, 69
 Perfect information, 32
 Periodic data-processing, 262 ff
 Perpetual auditing, 114
 Perpetual data-processing, 262 ff
 Pilferage costs, 11
 Planning, 4
 Planning horizon, 20, 88, 136 f
 Plan of book, 18
 Plant:
 capacity, 4, 5, 6, 19
 expansion, 6
 new, 4
P-model, simulation, 236 f, 239 ff, 247 f
 Potential savings percent, 180
 Precautionary motive, 17
 Present worth, 35
 savings, 173
 Price changes, 84

Prices, fluctuating, 143
 Probability distribution:
 continuous, 40
 discrete, 40
 Process flow charts, 216
 Procurement, 5
 Procurement costs, 9, 10
 Production, continuous, 10
 Profit, 23
 Programming, dynamic (*see* Dynamic programming)
 Protection, level, 164
 P-system, 114 f, 116, 125 ff, 160 ff

Q

Q-model, simulation, 224, 225, 246 f
 Q-model, updating routine, 327
 Q-system, 114, 116, 121 ff, 155 ff
 Quality control standards, 19
 Quality control, statistical, 319
 Quantity discounts, 84 ff
 Queuing equations, 278
 Queuing model, 146 ff, 236
 Queuing model, simulation, 242 ff
 Queuing theory, 148, 275, 288

R

Raiffa, Howard, 51
 Railroads, 4
 Ramp input, 301
 Rationality, bounded, 210, 257
 Receipt transaction, flow diagram, 286
 Redundancy, 322
 Regression analysis, 14
 Regret criterion, 51
 Regulated inventory systems, 296 ff
 Regulation, error, 297
 Regulator, 219
 Repetitive decisions, 6
 Replacement, jet engine, 242
 Research programs, 14
 Reserve stock, 113
 Reserve stock costs, controlling, 52
 Retailer, 19
 Retirement of executives, 14
 Review cost, 106
 Risk, 8, 16, 17, 18, 56
 Risk, cost of, 31 ff

Robinson, G., 334
 Rounding off months between orders, 81 ff
 Routine, data-processing, 263
 Routine, updating, 235, 325 ff
 Q-model, 327
 Ruin, gambler's, 201 ff
 Ruin, probability of, 203, 204
 Rule box, 218

S

Salvage value, 11, 20, 43
 Sample for evaluation study, 196 ff
 Sampling items, 197 f
 Savage, Leonard J., 51 f, 58
 Savage criterion, 146
 Savings, expected, 184, 185
 Savings, potential, 180
 Savings, present worth, 173
 Savings, probability distribution of, 202
 Savings vs. error rate, 177
 Scarf, H., 17
 Scheduling production, 20
 Schlaifer, Robert, 33
 Schoolroom capacity, 19
 Seasonal demand, 113
 Self-regulated inventory systems, 309 ff
 Self-regulation, total, 297
 Self-repairing machine, 236
 Self-supplying, 5, 9, 10, 12, 20
 Selling policy, optimal, 142 ff
 Sensitivity, 14
 Sensitivity analysis, 177 f
 Sequencing flow, 216
 Sequencing production, 20
 Service-time distribution, 277
 Servo-control model, 297
 Servomechanisms, 297
 Set-up cost, 9, 10, 19
 Set-ups, 79, 215
 Shannon, C., 321
 Shewhart, Walter A., 317
 Shortage of working capital, 95 f
 Siegel, Sidney, 61
 Simon, Herbert A., 14, 210, 298
 Simulation, 18, 218, 257 f, 283 f
 Simulation and system designs, 221 ff
 Simulation, inventory queuing model, 242 ff, 245
 Simulation, many item, 248

Simulation, *Q*-model, 224, 227 ff, 234
 Simulation, *Q*-model analog of queuing model, 246 f
 Simulation, *P*-model, 236 ff
 Simulation, *P*-model analog of queuing model, 247 f
 Simulation, *P*-model with variable lead time, 239 ff
 Size of inventory investment, 176 ff
 Skewness, 60, 61, 69, 154
 Smoothing, exponential, 320
 Spare parts, 33
 Special cases, 3, 5, 7
 Specialization, 281 ff
 Speculative motive, 17
 Square-root algorithm, 91
 (s, S)—policies, 115 f
 Stability, 252, 317
 Stable distributions, 317
 Starr, Martin K., 16, 201
 Static inventory problems, 6, 11, 18
 Stationarity, 318
 Stationary demand distribution, 117, 119
 Statistical quality control, 319
 Steel girders, 5
 Stochastic process, 131
 Stock-keeping units, 93
 Stocks, liquid, 36
 Stockage costs, 10
 Storage costs, 11
 Storage of intangibles, 14
 Strategies, 21, 22, 33
 Structure of analysis, 7 f
 Structure of inventory problems, 4 f
 Study, breadth, 175
 Study, depth, 175
 Study design, 18
 Study, evaluation, 171
 Study funds, exhaustion of, 201 ff
 Study, inventory, 172 ff
 Subjective probabilities, 55 ff
 Subjectivist school, 55
 Sunk costs, 4, 9, 10
 Supermarket inventory problems, 3, 5, 6, 112
 Super-*P* model, 310
 Switching matrices, 130 ff
 Symmetry, 61, 66, 67, 69, 73, 155, 163
 Synchronization of commodity flows, 17

Systemic costs, 12 f, 18, 109, 110, 115, 116, 124, 130, 171, 205, 295, 324
 Systems analysis, 18, 209 f
 System designs and simulation, 221 ff

T

Tangibility of stock, 4, 14, 15
 Taxicabs, 5
 Tchebycheff's inequality, 62 ff, 66 f, 156 f
 Telemetering, 290
 Telephone company, 45
 Time lag, 113, 114
 Time, lag (*see* Lead time)
 Total self-regulation, 297
 Transactional motive, 17
 Transients, 252
 Transition probabilities, 132
 Transitions, 131
 Trend input, 301, 303
 Tri-divisional classification, 181
 Two-bin system, 114

U

Uncertainty, 5, 6, 8, 18, 320
 Uncertainty and risk, 49 f
 Uncertainty, decision criteria, 50 ff, 53 ff
 Undamped oscillation, 314
 Underloaded data-processing systems, 269 ff
 Understock cost, 113
 Unimodality, 61, 69 f, 73, 155, 163
 Union regulations, 10
 Updating, 324 ff
 Updating controls, 316 ff
 Updating routine, 235, 325 f, 327
 Usage, variance, 181 ff
 Uspensky, J. V., 201
 Utility, 16, 25, 53
 Utilization factor, 242

V

Value, economic, 4
 Van Horn, 234
 Variable lags, 146 ff
 Variance, dollar usage, 186

Variance, forecast error, 194 ff
Variance of demand, 192
Variance of usage, 181 ff
Variety, 289, 321
Variety of commodities, 3

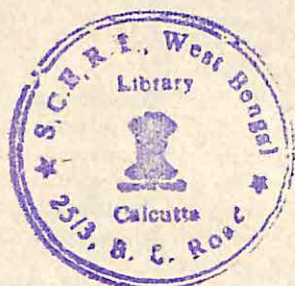
W

Wagner, Harvey M., 86
Wald, Abraham, 51
Wald criterion, 53
Warehouse, central, 162 ff, 213
Warehouses, multiple, 110
Weapon systems, 210

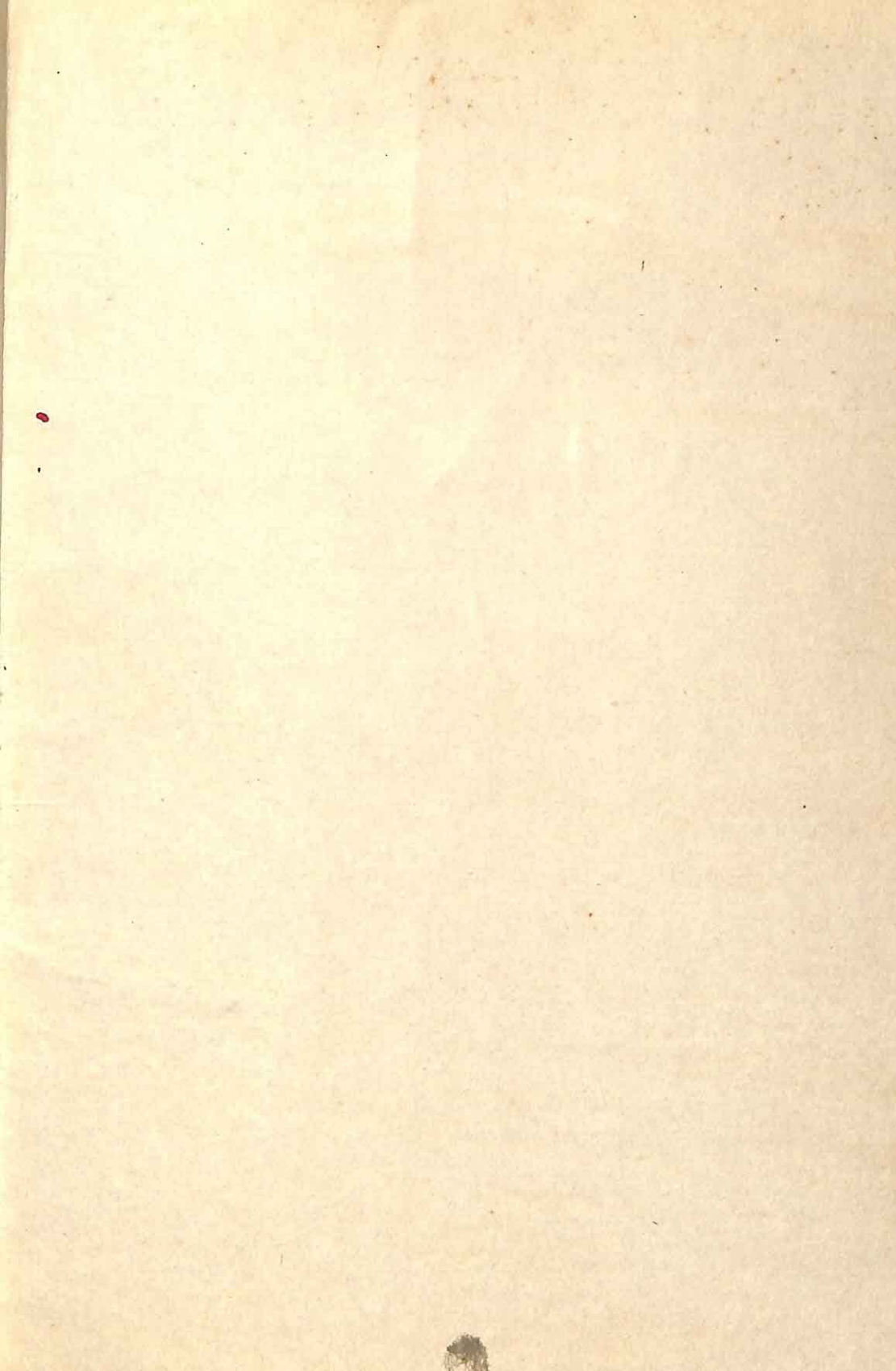
Whitin, T. M., 84, 86
Whittaker, Edmund, 334
Wolfowitz, Jacob, 115
Work force, size of, 45 f
Working capital, inventory problem of,
4
Working capital, shortage, 95 f
Work-pool, 289
Work sampling, 261

Z

z-transform, 304 ff
table of, 305







You'll be interested in
these other

**Eastern
Economy
Editions**

PRODUCTION AND INVENTORY CONTROL : Principles and Techniques, 2nd Ed.

by **George W. Plossl**, *G.W. Plossl and Co., Inc.*

This highly successful, thoroughly updated volume presents a *practical* approach to production control, focusing on basic principles and complete coverage of techniques rather than upon mathematical derivations.

Appropriate for courses in inventory and production control, production management and manufacturing control systems in departments of Industrial Engineering and Business Administration at the undergraduate/postgraduate level.

448 pages

MOTION AND TIME STUDY : IMPROVING PRODUCTIVITY, 5th Ed.

by **Marvin E. Mundel**, *M. E. Mundel & Associates, Maryland.*

A systematic, practical and logical treatment of motion and time study. Applied to a wide variety of activities including government organizations, agricultural and office work. An array of charts, graphs, illustrations, tables, and bibliographies make it a valuable text, reference and handbook for students and professionals.

**Prentice-Hall of India
Private Limited
New Delhi**

ISBN-0-87692-184-5

